

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: _____ Examiner #: _____ Date: _____
 Art Unit: _____ Phone Number 30 _____ Serial Number: _____
 Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

**For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

STAFF USE ONLY**Type of Search****Vendors and cost where applicable**

Searcher: B. D. Schneider NA Sequence (#) 1 STN _____
 Searcher Phone #: 308-4292 AA Sequence (#) 1 Dialog _____
 Searcher Location: CMC 6.103 Structure (#) _____ Questel/Orbit _____
 Date Searcher Picked Up: _____ Bibliographic _____ Dr. Link _____
 Date Completed: 5/22 Litigation _____ Lexis/Nexis _____
 Searcher Prep & Review Time: 10 Fulltext _____ Sequence Systems Cambridge
 Clerical Prep Time: _____ Patent Family _____ WWW/Internet _____
 Online Time: 6.44 Other _____ Other (specify) _____

THIS PAGE BLANK (USPTO)

Thu May 22 09:39:18 2003

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 21, 2003, 22:05:28 ; Search time 39 Seconds
(without alignments)
1303.977 Million cell updates/sec

Title: US-09-701-395A-23
Perfect score: 2743
Sequence: 1 MELLGVRLNLISSCPVWTFGT.....VRHLLSDPSGAVMVRAYLER 529

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_73.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1909.5	69.6	526	T07082	lycopene epsilon-cycl
2	749.5	27.3	500	S72505	lycopene beta-cycl
3	749.5	27.3	500	S66350	lycopene beta-cycl
4	737	26.9	500	S72506	lycopene beta-cycl
5	729	26.6	500	S66349	lycopene beta-cycl
6	705.5	25.7	498	S71511	capsanthin/capsoru
7	650.5	23.7	471	S51511	lycopene cyclase -
8	429.5	15.7	410	D75475	UbiH family protei
9	151.5	5.5	394	F81212	probable monooxyge
10	146.5	5.3	394	A81789	probable squalene
11	137.5	5.0	506	T07942	hypothetical prote
12	134.5	4.9	408	D70549	crty protein - Erw
13	125.5	4.6	382	C37802	lycopene cyclase -
14	125	4.6	386	S52981	bacteriochlorophyl
15	123	4.5	393	A75205	squalene monooxyge
16	116.5	4.2	408	H87193	probable FAB-link
17	116.5	4.2	408	H87193	NADH oxidase (H2O-
18	116.5	4.2	457	JC4541	hypothetical prote
19	115	4.2	704	AE2246	probable glutamate
20	113	4.1	476	D71200	probable lycopene
21	112.5	4.1	405	T37022	reverse gyrase (in
22	112.5	4.1	1613	T26448	probable squalene
23	112	4.1	518	T07940	2-octaprenyl-6-met
24	111.5	4.1	452	AE3596	hypothetical prote
25	111.5	4.1	456	DB2402	1-aspartate oxidas
26	111.5	4.1	550	B82074	hypothetical prote
27	111	4.0	517	T51363	probable squalene
28	110.5	4.0	381	H69096	bacteriochlorophyl
29	110.5	4.0	382	S52585	lycopene cyclase -

RESULT 1

T07082

lycopene epsilon-cyclase (EC 5.5.1.1) - tomato

C:Species: Lycopersicon esculentum (tomato)

C:Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 19-Apr-2002

C:Accession: T07082

R:Ronen, G.; Cohen, M.; Zamir, D.; Hirschberg, J.

Submitted to the EMBL Data Library, July 1997

A:Description: Regulation of expression of the gene for lycopene epsilon cyclase dur

A:Reference number: Z15905

A:Accession: T07082

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-526 <R0N>

A:Cross-references: EMBL:Y14387; NID:el285211; PIDN:CAA74745.1; PID:el285212

A:Experimental source: cultivar VF36; leaf

C:Genetics:

A:Gene: CrtL-e-1

A:Map position: 12

C:Function:

A:Description: converts lycopene to delta-carotene

A:Superfamily: tomato lycopene beta-cyclase

C:Keywords: Intramolecular lyase; intramolecular oxidoreductase; isomerase

Query Match 69.6%; Score 1909.5; DB 2; Length 526;
Best Local Similarity 68.5%; Pred. No. 8.8e-144;
Matches 368; Conservative 68; Mismatches 78; Indels 23; Gaps 6;

QY	1	MELLGVRLNLISSCPV-----WTFGTRNLSSK---LAYNHRVYSGSSCRVDFQVRADGG 50	QY	1	MECVGVQNVGAMAVLTPRLNRWGG--ELCQESIFLAY--EQYESK-----NSS 48
DB	1	SGSRSSVAYKEGFVDEEDFIRAGGSELLFVQMQOTKSMKEAKLADKLPPIPFSGSVMDL 110	DB	1	SGSRSSVAYKEGFVDEEDFIRAGGSELLFVQMQOTKSMKEAKLADKLPPIPFSGSVMDL 110
QY	51	SGSRSSVAYKEGFVDEEDFIRAGGSELLFVQMQOTKSMKEAKLADKLPPIPFSGSVMDL 110	QY	51	SGSRSSVAYKEGFVDEEDFIRAGGSELLFVQMQOTKSMKEAKLADKLPPIPFSGSVMDL 110
DB	49	SGSDSCVVDKEDFADEEDYIRAGGSQVLFVQMQKDDQSKLDELRLQISAGQTVLDL 108	DB	49	SGSDSCVVDKEDFADEEDYIRAGGSQVLFVQMQKDDQSKLDELRLQISAGQTVLDL 108
QY	111	VVIGCGPAGLSAFAAKLGLKVLIGLDPDLPFTNNYGVWEDEFKDLGLERCIEHAWKDTI 170	QY	111	VVIGCGPAGLSAFAAKLGLKVLIGLDPDLPFTNNYGVWEDEFKDLGLERCIEHAWKDTI 170
DB	109	VVIGCGPAGLSAFAAKLGLKVLIGLDPDLPFTNNYGVWEDEFKDLGLERCIEHAWKDTI 168	DB	109	VVIGCGPAGLSAFAAKLGLKVLIGLDPDLPFTNNYGVWEDEFKDLGLERCIEHAWKDTI 168
QY	171	VYLDNDAPVLLIGRAYGRVSRHLLHEELLKRCVSGSVYLDKSRVERITEAGDGHSLVVCVN 230	QY	171	VYLDNDAPVLLIGRAYGRVSRHLLHEELLKRCVSGSVYLDKSRVERITEAGDGHSLVVCVN 230
DB	169	VYLDNDAPVLLIGRAYGRVSRHLLHEELLKRCVSGSVYLDKSRVERITEAGDGHSLVVCVN 228	DB	169	VYLDNDAPVLLIGRAYGRVSRHLLHEELLKRCVSGSVYLDKSRVERITEAGDGHSLVVCVN 228
QY	231	EIFIFPCRLATVAGSAAKLLLEYVGGPRVCVQTAYGVEVEVENPYDNLAVFMDYDY 290	QY	231	EIFIFPCRLATVAGSAAKLLLEYVGGPRVCVQTAYGVEVEVENPYDNLAVFMDYDY 290
DB	229	DVVFICRFVTVAGSAAKLLLEYVGGPRVCVQTAYGVEVEVENPYDNLAVFMDYDY 288	DB	229	DVVFICRFVTVAGSAAKLLLEYVGGPRVCVQTAYGVEVEVENPYDNLAVFMDYDY 288
QY	291	MOOKLQCSSEETPTFLYVPMSPTRLFFFEETCLASDKAMPFDLLKKLSRLKTLGIQVT 350	QY	291	MOOKLQCSSEETPTFLYVPMSPTRLFFFEETCLASDKAMPFDLLKKLSRLKTLGIQVT 350
DB	289	LRHDAQSLEAKYPTFLYVPMSPTRLFFFEETCLASDKAMPFDLLKKLSRLKTLGIQVT 348	DB	289	LRHDAQSLEAKYPTFLYVPMSPTRLFFFEETCLASDKAMPFDLLKKLSRLKTLGIQVT 348

ALIGNMENTS

30	109.5	4.0	391	2	D64366	hypothetical prote
31	108.5	4.0	291	2	H97090	thioedoxin reduct
32	107	3.9	1206	2	AG2140	hypothetical prote
33	106.5	3.9	372	2	A75185	cell division prot
34	106	3.9	431	2	H97805	ubiH protein (EC 1
35	106	3.9	531	2	T52462	hypothetical prote
36	105.5	3.8	410	2	H69307	bacteriochlorophyl
37	105.5	3.8	414	1	S74699	probable 2-octapre
38	105	3.8	393	2	C71240	hypothetical prote
39	105	3.8	435	2	F89898	glucose-inhibited
40	104.5	3.8	302	2	G71660	hypothetical prote
41	104.5	3.8	430	2	T34627	probable electron
42	104.5	3.8	437	2	H95409	probable oxidoredu
43	104.5	3.8	516	2	T51364	probable squalene
44	104	3.8	659	2	H69421	heterodisulfide re
45	103.5	3.8	481	2	H97533	dihydrolipoamide d

[illegible]

RESULT 2
S72505
lycopene beta-cyclase (EC 5.5.1.-) - tomato
C:Species: Lycopersicon esculentum (tomato)
C:Date: 17-Feb-1998 #sequence_revision 26-Feb-1998 #text_change 28-Jul-2000
C:Accession: S72505
R:Pecker, I.; Gabbay, R.; Cunningham Jr., F.X.; Hirschberg, J.
submitted to the EMBL Data Library, April 1995
A:Description: Cloning and characterization of the lycopene cyclase gene from tomato.
A:Reference number: S72505
A:Accession: S72505
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-500 <PEC>
A:Cross-references: EMBL:X86452; NID:g1006672; PID:g1006673
C:Superfamily: tomato lycopene beta-cyclase
C:Keywords: tetramolecular lyase; isomerase

Query Match	27.38;	Score	749.5;	DB	2;	Length	500;
Best Local Similarity	36.6%;	Pred. No.	1.5e-51;				
Matches	169;	Conservative	92;	Mismatches	168;	Indels	33;
Gaps							
QY	73	GSSELLFVQMOTKSMERQAKLADLPPI-PFGESVMDLVVIGCGPAGLSAAEAAKLGL	131				
Db	52	GSSSALLEIVPET----	107				
QY	132	KVGLLIGPD--LPFTNNYGWDEDFKDLGLERCIEHAWKDTIVYLDNDAPVLIGRAYGRVS	189				
Db	108	SVCSIDPNPKLPNNYGVWDEFAMDLDCDWTSGAAVYIDNTAKDLHRYGRVN	167				
QY	190	RHLHHEELKRCVSGSVYLDSEKRITTEAGDHSVLVCENEIFPCRLATVSGAASGK	249				
Db	168	RQKASKMMQKIMNGVFKHQAIVKIWE--ESKSMLCNDGITTQATVVDATG-FGRS	225				
QY	250	LLEYEVG-GPRVCVQTAGEVEVENNPDPNLVFMRYD-YMQOKLQCSB--EYPTF	305				
Db	226	LVQYDKPNPG--YQVAYGILAEVEHPDVKNMFMFMWRDMSHLKNTDLKERNRIPTF	283				
QY	306	LVVPMSPTRLFETCTASKAMPEDLLKRLMSRLKTGLQTVKTVYEEESYIPVGS	365				
Db	284	LYAMPSSNRILFETLSVLRPGLRIDDIQERVARNLHGIKVKISIEDEHCLIPMGPG	343				
QY	366	LPNTEQKNLAFGAASWVHPATGYSVVRSLSEAPKYASVIAKLLQDNSAYVVGSSAV	425				
Db	344	LPVLPPRVVIGGTAGWHPSTGYMVARTAAAPVYANAIIQYLGERS-----HSGN	396				
QY	426	NTSMQAWSSLPKRRKQRAFFLGLLELIVQLDIEATRFTFFFLPTMMWGMGLSSSL	485				
Db	397	ELSTAVWKDLWPIERRRQREFFCFGMDDLLKLDLPATRFFEDAFDLEPRYWHGFLSSRL	456				
QY	486	SSSDLVLFMYMFVLAPNSMR-----MSLVRRHLSD	516				
Db	457	FLPELVFGLSLFSPHASNTRSRFIMTKGTVPVLVNNLNNLQD	498				

RESULT 3
S66350
lycopene beta-cyclase (EC 5.5.1.-) - tomato
C;Species: Lycopersicon esculentum (tomato)
C;Date: 14-Feb-1997 #sequence revision 13-Mar-1997 #text change 28-Jul-2000

C:Accession: S66350
R:Pecker, I.; Gabbay, R.; Cunningham Jr., F.X.; Hirschberg, J.
Plant Mol. Biol. 30, 807-819, 1996
A:Title: Cloning and characterization of the cDNA for lycopene beta-cyclase from tomato
A:Reference number: S66349; MUID:96194462; PMID:8624411
A:Accession: S66350
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-500 <PEC>
A:Cross-references: EMBL:X86452
C:Superfamily: tomato lycopene beta-cyclase
C:Keywords: intramolecular lyase; isomerase

	Query Match	27.38;	Score 749.5;	DB 2;	Length 500;
	Best Local Similarity	36.6%;	Pred. No. 1.5e-51;		
	Matches 169;	Conservative	92;	Mismatches 168;	Indels 33; Gaps 11;
Qy	73	GSELLFVQMOQTKSMEQAKLADKLPII-PFGESVMDLVVIGCGPAGLSLAAPAAKLG	131		
Db	52	GSSSALLELVET---KKENLDFELPMYDFSGVVVDLAVVGGGPAGLVAQVSEAGL	107		
Qy	132	KVGLIGPD--LPFTNNYGWDEDFKDLGERLCIEHAWKDTIYVLDNDAPVLIGRAYGRVS	189		
Db	108	SVCSIDPNPKLIIPNNYGWVDEFEAMDLLCLDATWSGAAYIIDNTAKDLHRPYGRVN	167		
Qy	190	RHLLEELLKRCVESGVSLDSKYVERITEAGDGHSLVGCENEIFPCRLATVGAASGK	249		
Db	168	RQLKSKMQMKQIMNGVAFHQKVJKVVIHE-ESKMLICNDGITIQTAVVLDATG-FSKS	225		
Qy	250	LLEYEVG-GPRVCQVQATYGEVEYENNNYPDNLVFMFYRD-YMOOKLKQCSB--EYPTTF	305		
Db	226	LVOYDKPNPG--YQVAYGILAEVEHPDVKNVKVMFMDWRSHLKNNTDLKERNSRIFTE	283		
Qy	306	LYVPMSPTRLEFEETCLASKDAMPDILLKRLMSRLKTGLGIQVTKVVEEWSYIPVGGS	365		
Db	284	LYAMPSSNRIFLEETSIVARGLRIDDIQERVARLNLHGILKVKASIEDEHCLIPMGPP	343		
Qy	366	LPNTQKNLAFGAASVNHVPATGYSVYRSLSSAPKASYAVIAKLQDNISAYVYVQSSAY	425		
Db	344	LPVLPVQVVYIGGTAGVHPSTGYMVVARTLAAAPVAVNAIIQYLGSKS-----HSGN	396		
Qy	426	NTSMQAWSSIWPKERKORAFFLGLELIVQIDIEATFTFFTRFLPTMMWVFLGSSL	485		
Db	397	EUSTAVWKDLVPIERRQREFFCFGMIDILKLDLPATRFPPDAFPDLEPRYWHGFLSSRL	456		
Qy	486	SSFDLVLSMYIAFLVAPNSMR-----MSLVRHLLSD	516		
Db	457	ETDFELVEGSLFSHASNTSRFLMTKGTVPVLYNNMLNQD	498		

RESULT 4
S72506
lycopene beta-cyclase (EC 5.5.1.-) - common tobacco
C.Species: *Nicotiana tabacum* (common tobacco)
C.Date: 17-Feb-1998 #sequence_revision 26-Feb-1998 #text_change 28-Jul-2000
C.Accession: S72506
R.Pecker, I.; Gabbay, R.; Cunningham Jr., F.X.; Hirschberg, J.
submitted to the EMBL Data Library, September 1994
A.Description: Cloning and characterization of the lycopene cyclase gene fr
A.Reference number: S72506
A.accession: S72506
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-500 <PEC>
A.Cross-references: EMBL:X8187; NID:g1006699; PID:g1006690
C.Superfamily: tomato lycopene beta-cyclase
C.keywords: intramolecular lvase; isomerase

Query Match 26.9%; Score 737; DB 2; Length 500;
Best Local Similarity 33.4%; Pred. No. 1.5e-50;
Matches 174; Conservative 111; Mismatches 180; Indels 56; Gaps 15;
QV 14 PWTFTGRNLSKKLAIYHRYGSSCRVDQVRADGGSSRSVAYEGFVDEEDFIKAG 73

Db 16 PVHGFVSK--ASSFNSVPRHFGSR-----KICENWKGV-----CVKAK 53
 QY 74 GSELL-FVQMOOTKSMEOAKLADKLPIPPGCVSMDLVVWICGGPAGLSLAAEAAGLGLK 132
 Db 54 SSALLELPETKKNLDFELPMYD-----PSKGLVVDLAVVGGGPGAGLAVAAQVSEAGLS 108
 QY 133 VGLI--GPDLPFTNNYGVWEDEFKDLGLERCIEHAKDTIVYLDNDAPVLIGRAYGRYSR 190
 Db 109 VVSDPSPKLIWPNNGYVWVDEFEAMDLLDCLDATWSTGVYIDNTTKDLDRPYGRVNR 168
 QY 191 HLLHEELLKRCVSGSYLDSKVRITERAGDGHSLVWCNEIFIPCRLATVASGAASGLK 250
 Db 169 KQLASKMOKCILNGVFXHAKVIVHE-EAKSMLICNDGVTIQTAVVLDATG-FSRCL 226
 QY 251 LEYEVGGP-RVCVOTAYGVEVENNPYDNLVPMFYRD-YMOOKLOCSE--EYPTFL 306
 Db 227 VOYD--FYPKPGYQVAGIILAEVEEHFDTSKVIMDMWRDHLGNMELKERNKVPFTFL 284
 QY 307 YVPMSPTRLFFETCLASKDAMPFDLLKRLKLSRLKTLGIQVTKVYEEESYIPVGGSL 366
 Db 285 YAMPFSSNKIFLETSVLARPGLRMDDDIQRVAVRLNHLGKIVKSIEDEHCVIPMGSL 344
 QY 367 PNTQKKNLGAASVMPATGYSVRSLSSEAPKYASVIKILKODNSAYVVSQSSAVN 426
 Db 345 PVIPQVVGTTGGTAGLVHPSTGYMARTLAAAPVYVANAIIHYLGSEKD--LLGNLSAA- 401
 QY 427 ISMAQWSLWPKERKORAFPLFGLLEIVOLDIEATRTFFRFLPTWMMWFGSL 486
 Db 402 ----VWKDLWPIERRRQREFFCGMDILLKLDLPATRRFFDAFFDLEPRYWHGFLSSRLY 457
 QY 487 SFDLVFSMYMFLVAPNSMR-----MSLVRHLLSD 516
 Db 458 LPELIFFGLSLFRASNTSRIEIMTKGTLPVNMNLLQD 498

RESULT 5

S66349

lycopene beta-cyclase (EC 5.5.1.1) - common tobacco

C:Species: Nicotiana tabacum (common tobacco)

C:Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 28-Jul-2000

C:Accession: S66349

R:Pecker, I.; Gabbay, R.; Cunningham Jr., F. X.; Hirschberg, J.

Plant Mol. Biol. 30, 807-819, 1996

A:Title: Cloning and characterization of the cDNA for lycopene beta-cyclase from tomato

A:Reference number: S66349; MUID:96194462; PMID:8624411

A:Accession: S66349

A>Status: preliminary; nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-500 <PEC>

A:Cross-references: EMBL:X81787

C:Superfamily: tomato lycopene beta-cyclase

C:Keywords: intramolecular lyase; isomerase

Query Match 26.6%; Score 729; DB 2; Length 500;

Best Local Similarity 35.1%; Pred. No. 6.3e-50;

Matches 163; Conservative 103; Mismatches 165; Indels 34; Gaps 12;

QY 70 IRAGSGELL-FVQMOOTKSMEOAKLADKLPIPPGCVSMDLVVWICGGPAGLSLAAEAAGL 128
 Db 50 VRAKSSALLELPETKKNLDFELPMYD-----PSKGLVVDLAVVGGGPGAGLAVAAQVSE 104
 QY 129 LGKLVGLI--GPDLPFTNNYGVWEDEFKDLGLERCIEHAKDTIVYLDNDAPVLIGRAYG 186
 Db 105 AGLSVYSIDPSKPLIWPNNYGVWVDEFEAMDLLDCLDATWGSVVVYIDNTTKDLDRPYG 164
 QY 187 RYVSRHLLHEELLKRCVSGSYLDSKVRITERAGDGHSLVWCNEIFIPCRLATVASGAAG 246
 Db 165 RVNRKOLKSKMOKCILNGVFXHAKVIVHE-EAKSMLICNDGVTIQTAVVLDATG-F 222
 QY 247 SGKLLLEYEVGGP-RVCVOTAYGVEVENNPYDNLVPMFYRD-YMOOKLOCSE--EY 302
 Db 223 SRCLVQYD--KPYKPGYQVAGIILAEVEEHFDTSKVIMDMWRDHLGNMELKERNKRV 280

QY 303 PTFLYVMPSPTRLFFETCLASKDAMPFDLLKRLKLSRLKTLGIQVTKVYEEESYIPV 362
 Db 281 PTFLYVMPSPTRLFFETCLASKDAMPFDLLKRLKLSRLKTLGIQVTKVYEEESYIPV 340
 QY 363 GGSLLPNTQKKNLGAASVMPATGYSVRSLSSEAPKYASVIKILKODNSAYVVSQSS 422
 Db 341 GGSLLPNTQKKNLGAASVMPATGYSVRSLSSEAPKYASVIKILKODNSAYVVSQSS 398
 QY 423 SAVNISMAQWSLWPKERKORAFPLFGLLEIVOLDIEATRTFFRFLPTWMMWFGSL 482
 Db 399 SAA-----VWKDLWPIERRRQREFFCGMDILLKLDLPATRRFFDAFFDLEPRYWHGFLS 453
 QY 483' SLSLSDFLVFSMYMFLVAPNSMR-----MSLVRHLLSD 516
 Db 454 SRLYLPELIFFGLSLFRASNTSRIEIMTKGTLPVNMNLLQD 498

RESULT 6

S71511

capsanthin/capsorubin synthase (EC 5.5.1.1) - pepper

N:Alternate names: chromoplast-specific protein

C:Species: Capsicum annuum (pepper)

C:Date: 09-Dec-1997 #sequence_revision 09-Dec-1997 #text_change 28-Jul-2000

C:Accession: S71511; JC2141

R:Bouvier, F.; Huquey, P.; d'Harlingue, A.; Kuntz, M.; Camara, B.

Plant J. 6, 45-54, 1994

A:Title: Xanthophyll biosynthesis in chromoplasts: Isolation and molecular cloning

A:Reference number: S71511; MUID:95004653; PMID:7920703

A:Accession: S71511

A:Molecule type: mRNA

A:Residues: 1-498 <BOU>

A:Cross-references: EMBL:X76165; NID:9522119; PIDN:CAA53759.1; PID:g522120

A:Experimental source: cv. Lamuyo; developmental stage ripening fruits

R:Deruere, J.; Bouvier, F.; Steppuhn, J.; Klein, A.; Camara, B.; Kuntz, M.

Biochem. Biophys. Res. Commun. 199, 1144-1150, 1994

A:Title: Structure and expression of two plant genes encoding chromoplast-specific

A:Reference number: JC2140; MUID:94197697; PMID:8147854

A:Accession: JC2141

A:Molecule type: DNA

A:Residues: 1-498 <DER>

A:Cross-references: EMBL:X77289; NID:9468747; PIDN:CAA54495.1; PID:g468748

R:Deruere, J.; Bouvier, F.; Steppuhn, J.; Klein, A.; Camara, B.; Kuntz, M.

Biochem. Biophys. Res. Commun. 201, 486, 1994

A:Reference number: A55775

A:Contents: annotation; erratum

A:Note: the legends of the nucleotide sequence figures were reversed in publication

C:Genetics:

A:Gene: CCS

C:Complex: monomer

C:Function:

A:Description: catalyses the conversion of the ubiquitous 5,6-epoxycartenoids, anthra

A:Pathway: carotenoid biosynthesis

C:Superfamily: tomato lycopene beta-cyclase

C:Keywords: chromoplast; intramolecular lyase; isomerase

Query Match 25.7%; Score 705.5; DB 2; Length 498;

Best Local Similarity 38.6%; Pred. No. 4.7e-48;

Matches 158; Conservative 79; Mismatches 153; Indels 19; Gaps 9;

QY 109 DLVYVIGCGPAGLSLAAEAAGLGLVGLIGDLPFT---NNYGVWEDEFKDLGLERCIEHA 165
 Db 83 DVIIIGPAGLRLAEQVSKYIKVCCVDFPS-PLSMPPNNYGVWVDEFEKLGLDECDLHK 141
 QY 166 WKDTIVYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVSGSYLDSKVRITERAGDGHSL 225
 Db 142 WPVSCVHISHDKYKYLDRPGRVSRKLLKLLNSCVENRVKFKAKVKKHBE-EPESS 200
 QY 226 VVCNEIFIPCRLATVASGAAGSKLLEYEVGGP-RVCVOTAYGVEVENNPYDNLVPMFV 284
 Db 201 IVCDGGRKISGLTVDASGVAS-DFEYD--KPRNHGVQVAGIILAEVNDHNPDLDDKML 257

QY 285 MDYRDY---MOOKLQCESEYPTFLYVMPSPTRLFFETCTCLASKDAMPFDLLKRLMSR 341
 Db 258 MDWRSHLGNPEYLRVKNKTEPTFLYAMPFDRLNLFLEETSLVSRPMLSYMEVRRVVR 317
 QY 342 LKTLGIQVTKVVEEWSYIPVGGSLPNTQKNLAFGAASVMVHPATGYSVVRSLSEAPKY 401
 Db 318 LRHLGIKVRSLVEEKVCVITMGGLPRIPQNVMAIGGTSIVHPSSGGYVARSMALAPVL 377
 QY 402 ASVIAKILQDNSAYVVSQSSAVNISMQAWSSLPKPKRQRAFFLGLLEIVOLDIEA 461
 Db 378 AEAIWESL---GSTRMIRGS---QLYHRVWNLWPSDRRRVRECYCFGMETLLKLDLEG 430
 QY 462 TRTFRTFFRLPTMMWGLFGSSLSFDLVFLFSFMFLVAPNSMRMSLV 510
 Db 431 TRRLDAFFDVPKYWHGFLSRSLSVKELAVLSLFLGHASNLARLDIV 479

RESULT 7
 S51511
 capsanthin/capsorubin synthase (EC 5.5.1.1) - pepper
 C:Species: Capsicum annuum (pepper)
 C:Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 28-Jul-2000
 C:Accession: S51511
 R:Houlne, G.; Schantz, M.L.; Meyer, B.; Pozueta-Romero, J.; Schantz, R.
 Curr. Genet. 26, 524-527, 1994
 A:Title: A chromoplast-specific protein in Capsicum annum: characterization and expression
 A:Reference number: S51511; MUID:95179811; PMID:7874747
 A:Accession: S51511
 A>Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-471 <HOU>
 A:Cross-references: GB:X78030; NID:g840728; PIDN:CAA54961.1; PID:g840729
 C:Superfamily: tomato lycopene beta-cyclase
 C:Keywords: intramolecular lyase; isomerase

Query Match 23.7%; Score 650.5; DB 2; Length 471;
 Best Local Similarity 38.2%; Pred. No. 1e-43;
 Matches 147; Conservative 75; Mismatches 144; Indels 19; Gaps 9;
 QY 109 DLWVIGCGPAGLSLAAEAAKLGKVLIGLPDLPFT---NNYGVWDEFEKDLGLERCIHA 165
 Db 83 DVIITGTGAGLRLAEQYSKYIKVCCVDPSP-PLSMWPNYGVWDEFEKLGLEDCLDHK 141
 QY 166 WKDTIVYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVSGVSYLDSKVERITEAGDGHSL 225
 Db 142 WPVSCVHISDHKTLYLDRPYGRVSRKLLKLLNSCVENRVRFYAKVLRKXHE-EFESS 200
 QY 226 VVCENEIFPCRLATVAVSAGSKLLEYEVGGR-VCVQTAYGVEVEVNNYDPLMYF 284
 Db 201 IVCDDGRKISGLIVDASGYAS-DFEYD--KPRNHGYQVAHGILAEVDNHPFDLDMML 257
 QY 285 MDYRDY---MOOKLQCESEYPTFLYVMPSPTRLFFETCTCLASKDAMPFDLLKRLMSR 341
 Db 258 MDWRSHLGNPEYLRVKNKTEPTFLYAMPFDRLNLFLEETSLVSRPMLSYMEVRRVVR 317
 QY 342 LKTLGIQVTKVVEEWSYIPVGGSLPNTQKNLAFGAASVMVHPATGYSVVRSLSEAPKY 401
 Db 318 LRHLGIKVRSLVEEKVCVITMGGLPRIPQNVMAIGGTSIVHPSSGGYVARSMALAP-- 375
 QY 402 ASVIAKILQDNSAYVVSQSSAVNISMQAWSSLPKPKRQRAFFLGLLEIVOLDIEA 461
 Db 376 -VLLRPVIVSLGSTRMIRGS---QLYHRVWNLWLRHRRRRVRECYCFGMETLLKLDLEG 430
 QY 462 TRTFRTFFRLPTMMWGLFGSSLS 486
 Db 431 TRRLDAFFDVPKYWHGFLSRSL 455

RESULT 8
 D75475
 lycopene cyclase - Deinococcus radiodurans (strain RL)
 C:Species: Deinococcus radiodurans

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
 C:Accession: D75475
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.
 M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Otterback, T.; Zalewski, C.;
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A:Reference number: A75250; MUID:20036896; PMID:10567266
 A:Accession: D75475
 A>Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-410 <WHI>
 A:Cross-references: GB:AE001934; GB:AE00513; NID:96458504; PIDN:AAF10377.1; PID:g645
 A:Experimental source: strain R1
 C:Genetics:
 A:Gene: DR0801
 A:Map position: 1

Query Match 15.7%; Score 429.5; DB 2; Length 410;
 Best Local Similarity 31.0%; Pred. No. 3.1e-26;
 Matches 135; Conservative 53; Mismatches 204; Indels 43; Gaps 13;
 QY 102 PFGESVMDLVVIGCGPAGLSLAAEAAKLGKVLIG--PDLPTNNYGVWDEFEKDLG-- 157
 Db 3 PPSPASSDVLVIGGPGSGTALSRAELARGLDVQQLAPHPRPFPATYGAW-----LGD 56
 QY 158 ---LERCIEHAWKDTIVYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVE--SGVS 212
 Db 57 PTWARGCAEQVWTDVRAY--TGPQPTSLGQPYA-----LLDNAALLTLRLGLADWTVEGA 110
 QY 213 VERITEAGDGHSLVVCENEIFPCRLATVAVSAGSKLLEYEVGGRVCVQTAYGVEVEV 272
 Db 111 ALHAERSGAGWTVYGAGGERW--QTRLVVDASG--HGALVSPVRFPGCAALQATYGVVARF 167
 QY 273 ENNPDYDNLVFMVMDYDMQKQKLCSEETFLYVMPSPTRLFFETCTCLASKDAMPFD 332
 Db 168 RRPVTPGSMWMDYRT-PAPELKRGE---ATFLYAMHLGGDRFYFEETSLIARAPTRA 223
 QY 333 LKRLKMSRLKVLGTQVTKVVEEWSYIPVGGSLPNTQKNLAFGAASVMVHPATGYSV 392
 Db 224 ELRRLRLARLSAQGTTPPHATESEEWAFPMNAQAP--APGVLAYAGAAAGRVHPVGFQVA 282
 QY 393 RSLSEAPKYASVIAKILQDNSAYVVSQSSAVNISMQAWSSLPKPKRQRAFFLGL 452
 Db 283 GALSADPQVATATATACQKDA-----AAAGWAALWSPERRAAREVHLLGVG 330
 QY 453 LIVQLDIEATRTFFRTFLPTMMWGLFGSSLSFDLVFLFSFMFLVAPNSMRMSLV 512
 Db 331 ALLGLERELPHFFGTFGLPREQWARFLHPDTPDAGTLARTMLRVFAQTGGVRLPLARA 390
 QY 513 LLSDP--SCAVMVR 525
 Db 391 ALAQPAASGRALAAA 405

RESULT 9
 F81212
 UbiH family protein NME0323 [imported] - Neisseria meningitidis (strain MC58 serogro
 C:Species: Neisseria meningitidis
 C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
 C:Accession: F81212
 R:Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen,
 Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.
 ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
 Science 287, 1809-1815, 2000
 A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.
 A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
 A:Reference number: A81000; MUID:20175755; PMID:10710307
 A:Accession: F81212
 A>Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-394 <TET>
 A:Cross-references: GB:AE002389; GB:AE002098; NID:g7225537; PIDN:AAF40768.1; PID:g72

us-09-701-395a-23.rpr

Thu May 22 09:39:18 2003

A:Experimental source: serogroup B, strain MC58

C:Genetics:

A:Gene: NMB0323

Query Match 5.5%; Score 151.5; DB 2; Length 394;
Best Local Similarity 21.7%; Pred. No. 0.00036;
Matches 84; Conservative 57; Mismatches 133; Indels 113; Gaps 17;

QY 109 DLVVICGPGAGLSAAEAKGLKVLII-----GPDLPFTN-----NYG 147
DB 6 DILVVGAGPAGLSFAAEAGSLKLVLIERSPLTVLQNPDPDGREIALTHFSREIMORLG 65
QY 148 VW-----EDEFKDLGLERCIEHAWKDTIVYLDNDAPVLIGR-----AYGRVS-- 189
DB 66 MWDKIPENE-----IYPLRDAKVLNGRSYDQLHFPOPTGEARPADC 107
QY 190 -RHLHEELLKRCVSGSVYLD-----KVERITEAGDGHSLVCENEIFIPCLATVA 242
DB 108 LGLYLSNNHNRRAAYEVVQSDNVLSITDVVKEVKTSDNEAQVILENGKILTARLLAA 167
QY 243 SGAASGKLEXYEVGPRVCQTAYGVEVEVNNPDYDNLVMDYDYMOKLQCESEY 302
DB 168 DSRFS-----QTRRLGIGSSDMHDIYRTMFVCR-----MKHTLSNQHTAY 207
QY 303 PTFLY--VMPMSPTRLFFETCL-----ASKDAMPFDL-----LKRKLMSRLKTLGIQV 349
DB 208 ECFHYGRTIALLPLEEHLNTVITVDTKINSVQNLSPLELAASVKQLKGRIGDMEL-V 266
QY 350 TKVYEEMSYIPVGGSLPNT--EOKNLAFGAASVHPATGYSVVRSISEAPKYASVIK 407
DB 267 SSIH-----HYPLVGMIAKRFYKRSALIGDAAGVHPVTAHGFGNGLSS-----ADILAK 317
QY 408 ILKQDNGAYVVGSSAVNISMOWS 434
DB 318 LITAEQ-----RGQDIGASSLLEKYSS 340

RESULT 10

A:Experimental source: serogroup A, strain Z2491

C:Genetics:

A:Gene: NMA2164

Query Match 5.3%; Score 146.5; DB 2; Length 394;
Best Local Similarity 21.5%; Pred. No. 0.00095;
Matches 83; Conservative 56; Mismatches 130; Indels 117; Gaps 16;

QY 109 DLVVICGPGAGLSAAEAKGLKVLII-----GPDLPFTN-----NYG 147
DB 6 DILVVGAGPAGLSFAAEAGSLKLVLIERSPLTVLQNPDPDGREIALTHFSREIMORLG 65
QY 148 VW-----EDEFKDLGLERCIEHAWKDTIVYLDNDAPVLIGR-----AYGRVS-- 189
DB 66 MWDKIPENE-----IYPLRDAKVLNGRSYDQLHFPOPTGEARPADC 107
QY 190 -RHLHEELLKRCVSGSVYLD-----KVERITEAGDGHSLVCENEIFIPCLATVA 242
DB 108 LGLYLSNNHNRRAAYEVVQSDNVLSITDVVKEVKTSDNEAQVILENGKILTARLLAA 167

RESULT 10

A:Experimental source: serogroup A, strain Z2491

C:Genetics:

A:Gene: NMA2164

Query Match 5.0%; Score 137.5; DB 2; Length 506;
Best Local Similarity 24.3%; Pred. No. 0.0071;
Matches 108; Conservative 52; Mismatches 145; Indels 139; Gaps 25;

QY 89 EKOAKLADKLPPFPGESVMDLVVIGCGPAGLSAAEAKGLKVLIGLIPDL--PFTNNY 146
DB 30 KKVAKLPDAATEVR-RDGDADTVIIVGAGVGSALAYALAKDGRVHVTERDMREFVR-- 85
QY 147 GWDEDF-----KDLGLERCIE-----HAWKDTIVYL-----DNDAPV-LI 181
DB 86 --MMGEFMQPGGRLLLSKLGLEDCELDQIATGLAVYKDGOKALVSPEDNDFFPEPT 223
QY 182 GRAY--GRVSRHL-----LHEELKRCVE-----SGVSYLDSKVERITEAGDGH 223
DB 144 GRAFYNGRFVQRLKOKASSLFTVQLEBGTVSLIEKGVIKGVYTKNS-----AGE-- 194
QY 224 SLVVCENEIFIPCLATVAGSAA-----GKLLYEYGGPRVCQTAYGVEVEVE 273
DB 195 -----ETTAFAP--LTVVCDGCYSNLRSSVNDNNAEVIYSYOVG-----YVSKNCQLE-- 239
QY 274 NNPDNLVMDYDYMOKLQCESEYPTFLVYMPMSPTRLFFETCLASDKAM----- 329
DB 240 -----DPEKLKIMSKEPSFTMLYQISSTDVRCVMEIFFPGNISGEMAVYLNKTMAPQV 295
QY 330 PFDLLKRLKMSRLKTLGIQV---TKVYEEMSYIPVGGSLPNTPEOKNLAFGAASVHP 385
DB 296 PPEL--RKIFLKGIDEGAQIKAMPTKRMEATLS-----EQGVIVIGDADFNRHP 343
QY 386 --ATGYSVW-----RSLSEAPKYASVIKILKQDNGAYV--SGQSSAVNIS 428
DB 344 AIASGMVVLSDILILRLQLPLNLSNDANKVSEVI-----KSFYVIRKPMSTAVNTL 396
QY 429 MQAWSSLW-----PKERKQRAF 446

RESULT 11

A:Experimental source: serogroup A, strain Z2491

C:Genetics:

A:Gene: NMA2164

Query Match 5.0%; Score 137.5; DB 2; Length 506;
Best Local Similarity 24.3%; Pred. No. 0.0071;
Matches 108; Conservative 52; Mismatches 145; Indels 139; Gaps 25;

QY 89 EKOAKLADKLPPFPGESVMDLVVIGCGPAGLSAAEAKGLKVLIGLIPDL--PFTNNY 146
DB 30 KKVAKLPDAATEVR-RDGDADTVIIVGAGVGSALAYALAKDGRVHVTERDMREFVR-- 85
QY 147 GWDEDF-----KDLGLERCIE-----HAWKDTIVYL-----DNDAPV-LI 181
DB 86 --MMGEFMQPGGRLLLSKLGLEDCELDQIATGLAVYKDGOKALVSPEDNDFFPEPT 223
QY 182 GRAY--GRVSRHL-----LHEELKRCVE-----SGVSYLDSKVERITEAGDGH 223
DB 144 GRAFYNGRFVQRLKOKASSLFTVQLEBGTVSLIEKGVIKGVYTKNS-----AGE-- 194
QY 224 SLVVCENEIFIPCLATVAGSAA-----GKLLYEYGGPRVCQTAYGVEVEVE 273
DB 195 -----ETTAFAP--LTVVCDGCYSNLRSSVNDNNAEVIYSYOVG-----YVSKNCQLE-- 239
QY 274 NNPDNLVMDYDYMOKLQCESEYPTFLVYMPMSPTRLFFETCLASDKAM----- 329
DB 240 -----DPEKLKIMSKEPSFTMLYQISSTDVRCVMEIFFPGNISGEMAVYLNKTMAPQV 295
QY 330 PFDLLKRLKMSRLKTLGIQV---TKVYEEMSYIPVGGSLPNTPEOKNLAFGAASVHP 385
DB 296 PPEL--RKIFLKGIDEGAQIKAMPTKRMEATLS-----EQGVIVIGDADFNRHP 343
QY 386 --ATGYSVW-----RSLSEAPKYASVIKILKQDNGAYV--SGQSSAVNIS 428
DB 344 AIASGMVVLSDILILRLQLPLNLSNDANKVSEVI-----KSFYVIRKPMSTAVNTL 396
QY 429 MQAWSSLW-----PKERKQRAF 446

Db 397 GNAFSQVLIASDEAKEAMROGCF 420

RESULT 12

D70549

Hypothetical protein Rv0561c - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: D70549

R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Rajandream, M.A.; Rogers, K.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Nature 393, 537-544, 1998
A: Authors: Scares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrall, B.G.
A: Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A: Reference number: A70500; MUID: 98295987; PMID: 9634230
A: Accession: D70549

A: Status: Preliminary; nucleic acid sequence not shown; translation not shown
A: Molecule type: DNA
A: Residues: 1-408 <COL>
A: Cross-references: GB:95558; GB:AL123456; NID:93261781; PIDN: CAB08972.1; PID: g2114023
A: Experimental source: strain H37RV
C: Genetics:
A: Gene: Rv0561c
C: Superfamily: fixC protein

Query Match 4.9%; Score 134.5; DB 2; Length 408;
Best Local Similarity 23.8%; Pred. No. 0.009;
Matches 91; Conservative 45; Mismatches 130; Indels 117; Gaps 20;

QY 109 DLVVGCGPAGLSAAEAKGLKVLGIGP-----DLPFTN-----NYGVW- 149

Db 8 DVVVVGAGPAGSAAWAAARAGRDVLVDITATFPRDKPCGDLPRVAELHQLGLKWL 67

QY 150 EDEFKDLGLERC-----IEHAWKDTIVLDNDAPVLIGRAYG-RVSRHLLHEELLKRCVE 203

Db 68 ADHIRHRLGRMSFGGEVEVDN-----PGSPFSGYSAVARLELDDRIKVAED 116

QY 204 SGVSYL-----DS--KVERITEAGDGHSLVVCENEIFIPCLATVAVSAGAS--GK 249

Db 117 TGARMLLGAKAVAVHDSRRVSLTLA-DGTE-----VGCQLIVADGARSPLGR 166

QY 250 LLEVEVGPRVCVOTAYGVEVE-VENNPY--DNLVFMVDYRDYMOOKLQCSEEEYPTFL 306

Db 167 KL-----GRKWHRETVYGVAVGYLSTAYSDDPWLTSHLELRSPDCAVL-----PGYG 214

QY 307 YVMPSPTRLFFETCTCLAKDAMPDLKRLKMSRLKTLGIGVTKVYEEWSY----- 359

Db 215 WIFPLNGEVNIGVGL-STSRPADLALRPLIS-----YYTDLRDEMGFTGQPRV 266

QY 360 ----IPVGSPLNTEQKN-LAFGAASVMVHPATGYSVRSLSSEAPKYASVIKILKQDNS 414

Db 267 SSALLPWGAVSGVAGSNMMLIGDAACVPLNNGEIDYGLE----- 308

QY 415 AYVVGQSSAVNISMQAWSSSLWP 437

Db 309 ----TGLAAELDSRLARLWP 327

RESULT 13

C37802

crty protein - Erwinia uredovora

C:Species: Erwinia uredovora

C>Date: 31-May-1991 #sequence_revision 31-May-1991 #text_change 20-Jun-2000

C:Accession: C37802

R: Misawa, N.; Nakagawa, M.; Kobayashi, K.; Yamano, S.; Izawa, Y.; Nakamura, K.; Harashim

J. Bacteriol. 172, 6704-6712, 1990

A: Title: Elucidation of the Erwinia uredovora carotenoid biosynthetic pathway by functio

A: Reference number: A37802; MUID: 91072214; PMID: 2254247

A: Accession: C37802

A: Status: Preliminary

A: Molecule type: DNA

A: Residues: 1-382 <MIS>

A: Cross-references: GB:D90087; NID: g216681; PIDN: BAA14126.1; PID: g216684
C: Superfamily: Erwinia dycopene cyclase

Query Match

Best Local Similarity 4.6%; Score 125.5; DB 2; Length 386;

Matches 79; Conservative 74; Mismatches 185; Indels 57; Gaps 15;

QY 109 DLVVGCGPAG--LSLAAEAAKGLKVLGIGPDLPTNNYGVWEDEFKDL-----GLER 160

Db 6 DLILVGGGLANGLIARLQRYPOLNLLIEAGQPGNGNHTWSFHEDDLTPGQHAWLAPL 62

QY 161 CIEHAWKDTIVLDNDAPVLIGRAYGVRSHLLHEELLKRCVSGVSYLDSKVERITEA-- 219

Db 65 LVVHWPDYQVRRPTRRKUNSGYFCTTSQRF--AEVLQR--QFGPHLMWDATAVEV--- 117

QY 220 GDGHSLLVVCENEIFIPCLATVAVSAGASGLKLEYEYGVGRVCVQYAYGEVEVENNPYDP 279

Db 118 -NAESVRLKKGQV-IGARAVIDGRGYAANSALS-----VGFOAFICQENRL-SHPHGL 167

QY 280 NLMVFMVDYRDYMOOKLQCSEEEYPTFLVMPSPTRLFFETCTCLAKDAMPDLKRLKLM 339

Db 168 SSPIMD-----ATVDQNGYRFVYSLPSLPTLLIEDTHYIDNATLDPECARONIC 219

QY 340 SRLKTLGIGVTKVYEEWSYIPV---GGSLEPTEQKNLA-FGAASVMVHPATGYSVYRS 395

Db 220 DYAAOQGWOLTLREEOGALPITLSGNADAFWQORPLACSGRLAGLHFTTGYSLPLAV 279

QY 396 SPAPKYASVIKILKQDNSAVVSGSSAVNISMQAWSSSLWPKRKRQRAFF-LFGLLELI 454

Db 280 AVADR-----LSALDVTFSASIHAIHTHAFARERWQOQGFMLNRMFL 322

QY 455 VOLDIEATRTFFRFLPTMMWGLSGSLSSFD 489

Db 323 LAGPADSRVWMORFVGLPEDLIAREYAGKLITLD 357

RESULT 14

S52981

lycopene cyclase - Erwinia herbicola

C:Species: Erwinia herbicola

C>Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 10-Dec-1999

C:Accession: S52981

R: Hundle, B.; Alberti, M.; Niveststein, V.; Beyer, P.; Kleinig, H.; Armstrong, G.A.;

Mol. Gen. Genet. 245, 406-416, 1994

A: Title: Functional assignment of Erwinia herbicola Eho 10 carotenoid genes expressed

A: Reference number: S52976; MUID: 95107236; PMID: 7808389

A: Accession: S52981

A: Status: preliminary

A: Molecule type: DNA

A: Residues: 1-386 <HUN>

A: Cross-references: EMBL:M87280; NID: g148404; PIDN: AAA64980.1; PID: g148411

C: Genetics:

A: Start codon: GTC

C: Superfamily: Erwinia dycopene cyclase

Query Match

Best Local Similarity 4.6%; Score 125; DB 2; Length 386;

Matches 86; Conservative 62; Mismatches 163; Indels 112; Gaps 18;

QY 109 DLVVGCGPAG--GL---SLAAEAAKGLKVLGIGPDLPTNNYGVWEDEF---KDLGLERC 161

Db 3 DLILVGGGLANGLIARLQRYPOLNLLIEAGQPGNGNHTWSFHEDDLTPGQHAWLAPL 62

QY 162 IEHAWKDTIVLDNDAPVLIGRAYGVRSHLLHEELLKRCVSGVSYLDSKVERITEA-- 219

Db 63 VAHAWPGYEVQFPD-----LRRRLARG--YYSITSERFAEALH 98

QY 220 ---GDGHSLLVVCENEIFIPCLATV-----ASGAA--SGKLEYEYGVGRVCVQYAYG 267

Db 99 QALGE-----NIWLNCVSEVLPSNVRVLANGALLAGAVIDGRGVYASSAMQYQ 149

QY 268 VEVEVE---NNPYDPLNLMVFMVDYRDYMOOKLQCSEEEYPTFLVMPSPTRLFFETCTLA 324

Job time : 44 secs

Db 150 LFLGQOMRLTQPHGLVPIILMDATVAQQGYR-----FVYTLPLSADTLIEDTRYA 201
QY 325 SKDAMPFDLLKRLKMSRLKTLGIQVTKVYEEWSYIPV--GGSL-----PNTQKNLA 375
Db 202 NYPQRDDNALRQVTVYAHSGWQLAQLEREETGCLPITLAGDIOALWADAPCVPRS--- 258
QY 376 FGAASMVHPATGYS-----VYRSLSEAPKYASVIKILKODNSAYVVGSSAYNISM 429
Db 259 -GMRAGLFHPTTGYSPLAVALADATADSPRLGSV-----PLYQLTR 299
QY 430 QAWSSLPKE---RKQRAFFLEGLLELIVOLDIEATRTFFRFLPTWMMWGLGSSLS 486
Db 300 QFAERHWRQGFRLNRLMLAGRE-----ENRWYMQRFYGLPEPTVERFYAGRLS 352
QY 487 SFD 489
Db 353 LFD 355

RESULT 15
A75205
bacteriochlorophyll synthase 43 kda chain or geranylgeranyl hydrogenase PAB0109 - Pyrococcus
C:Species: Pyrococcus abyssi
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
C:Accession: A75205
R:anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru
A:Reference number: A75001
A:Accession: A75205
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-393 <KAW>
A:Cross-references: GB:AJ248283; GB:AL096836; NID:g5457433; PIDN:CAB49088.1; PID:g545759
A:Experimental source: strain Orsay
C:Genetics:
A:Gene: bchP; PAB0109
C:Superfamily: hypothetical protein MJ0532

Query Match 4.5%; Score 123; DB 2; Length 393;
Best Local Similarity 22.3%; Pred No. 0.07;
Matches 86; Conservative 55; Mismatches 181; Indels 64; Gaps 15;

QY 109 DLVWICGPGAGLSAAEAKGLKVLIGP-----DLPTNNYGVWDEPKDLGL----ERC 161
Db 4 DVVVVSGVAGPTVARDVAKAGFSVLLVDKAAIGTPKQCAEGITIEVEFEKFDIPYDKRF 63
QY 162 IEHANKDTIVLDNAPVLIGRAYGRVSRHLHEELLKRCVESGVSYLDSKVERITEAGD 221
Db 64 INREYGAKIYSPSGYTAEL--RYKHVSGVILERKVFDRKMLAYYAKAGAEVLARTEVVD 121
QY 222 GHSLVVCEN-----EIFPCRLATVAGSAAS--GKLLLEYVGGPRVCVQTAY 266
Db 122 ----VIRNGKIVGRVRAKHEGEFLEAKVIVAADGVESTIARKAGINTYAPPEFDSGY 177
QY 267 GVEVEVNNPDNLMVMDYDMQKLOCSEEEYPT-FLXVMPMSPTLFLFEETCLAS 325
Db 178 EYEMLIIEG--YDPDLI-----HLFFGNEVAPRGVWIFPKDEDR---ANVGIGI 221
QY 326 KDAMPFDLLKRLKMSRLKTLGIQVTKVYEEWSYIPVGGSLPNTQKN-LAFGAASMVH 384
Db 222 ASDHP-QTAKYYLDKWLKENNIPMRKILEVNVGLVPVGGFVKELVKDNVILVVGDAARQVN 280
QY 385 PATGYSVVRSLSEAPKYASVIKILKODNSAYVVGSSAYNISMQAWSSLP-----K 438
Db 281 PVHGGGMEAKAASIAAKWIKALEENLELL-----KNYSEEWRTGPRMEKLLK 333
QY 439 ERKQRAFFLGLLEIVQL-----DIE 460
Db 334 LRKAMEKLDELDVFLVLSGTDL 359

Search completed: May 21, 2003, 22:12:01

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 21, 2003, 22:29:03 ; Search time 233 Seconds
(without alignments)
2997.959 Million cell updates/sec

Title: US-09-701-395A-23
Perfect score: 2743
Sequence: 1 MELLGVRNLISCPWTFGT.....VRHLLSDPSGAVMYRAYLER 529

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 828747 seqs, 660231138 residues
Total number of hits satisfying chosen parameters: 1657494

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DRV=xlh
-Q=/cgn2_1/USPTO.spool/US09701395/runat_15052003_165754_3440/app_query.fasta.1.711
-DB=PublishedApplications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LISP=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09701395 -CGN 1_1 135 -runat_15052003_165754_3440
-NCPU=6 -ICPU=3 -NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PublishedApplications_NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2743	100.0	1898	10	US-09-323-998D-22
2	2728	99.5	1848	10	US-09-323-998D-46
3	1959.5	71.4	1860	10	US-09-323-998D-1
4	1478.5	53.9	1378	10	US-09-323-998D-24

5	1394	50.8	1272	12	US-10-041-472-3	Sequence 3, Appli
6	1202	43.8	960	12	US-10-041-472-2	Sequence 2, Appli
7	739	26.9	1590	12	US-10-041-472-4	Sequence 4, Appli
8	173	6.3	1185	9	US-10-166-037-2	Sequence 2, Appli
9	129.5	4.7	1149	9	US-09-941-947A-29	Sequence 29, Appli
10	123	4.5	1593	9	US-09-938-842A-1578	Sequence 1578, Ap
11	115.5	4.2	1272	9	US-09-738-626-529	Sequence 529, App
12	113	4.1	1149	10	US-09-547-267-8	Sequence 8, Appli
13	113	4.1	8625	9	US-09-920-923-1	Sequence 1, Appli
14	113	4.1	11233	9	US-09-920-923-27	Sequence 27, Appli
15	111.5	4.1	536165	9	US-09-939-964-1	Sequence 1, Appli
16	111	4.0	1778	10	US-09-887-576-634	Sequence 634, App
17	106	3.9	6796	7	US-08-781-986A-2	Sequence 2, Appli
18	104.5	3.8	1551	9	US-09-938-842A-1094	Sequence 1094, Ap
19	103	3.8	2326	9	US-10-025-380-1103	Sequence 1103, Ap
20	103	3.8	2326	10	US-09-922-217-1103	Sequence 1103, Ap
21	102	3.7	2025	10	US-09-070-927A-439	Sequence 439, App
22	101	3.7	1308	10	US-09-815-242-6599	Sequence 6599, Ap
23	101	3.7	1999	10	US-09-917-800A-1706	Sequence 1706, Ap
24	99.5	3.6	1704	10	US-09-815-242-9465	Sequence 9465, Ap
25	99	3.6	1407	9	US-09-738-626-764	Sequence 764, App
26	99	3.6	2389	9	US-10-010-920-1	Sequence 1, Appli
27	99	3.6	2389	9	US-10-270-877-1	Sequence 1, Appli
28	99	3.6	2389	9	US-10-270-877-1	Sequence 1, Appli
29	99	3.6	2389	9	US-10-008-721-1	Sequence 1, Appli
30	98	3.6	549	9	US-09-991-936-256	Sequence 256, App
31	98	3.6	1179	10	US-09-815-242-7698	Sequence 7698, Ap
32	98	3.6	1303	10	US-09-815-242-4838	Sequence 4838, Ap
33	98	3.6	1308	10	US-09-815-242-8761	Sequence 8761, Ap
34	97.5	3.6	1551	9	US-09-938-842A-2698	Sequence 2698, Ap
35	97.5	3.6	3014	9	US-09-933-261-1	Sequence 1, Appli
36	97.5	3.6	3014	9	US-10-256-702-1	Sequence 1, Appli
37	97.5	3.6	3309400	9	US-09-738-626-1	Sequence 1, Appli
38	97	3.5	1385	9	US-10-005-530-40	Sequence 40, Appli
39	95.5	3.5	2196	10	US-09-917-800A-1469	Sequence 1469, Ap
40	95	3.5	1260	10	US-09-815-242-9700	Sequence 9700, Ap
41	95	3.5	2206	9	US-10-108-605-258	Sequence 258, App
42	95	3.5	3989	10	US-09-070-927A-359	Sequence 359, App
43	94.5	3.4	2493	9	US-09-938-842A-322	Sequence 322, App
44	93.5	3.4	2169	9	US-09-906-419-8	Sequence 8, Appli
45	93.5	3.4	2186	12	US-10-062-254-91	Sequence 91, Appli

ALIGNMENTS

RESULT 1
US-09-323-998D-22
; Sequence 22, Application US/09323998D
; Patent No. US20020102631A1
; GENERAL INFORMATION:
; APPLICANT: CUNNINGHAM JR., FRANCIS X.
; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 108172-09019
; CURRENT APPLICATION NUMBER: US/09/323,998D
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: 09/088,724
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 09/088,725
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 08/937,155
; PRIOR FILING DATE: 1997-09-25
; PRIOR APPLICATION NUMBER: 08/624,125
; PRIOR FILING DATE: 1996-03-29
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 1898
; TYPE: DNA
; ORGANISM: Adonis palaestina
US-09-323-998D-22

Alignment Scores:

Pred. No.: 0 Length: 1898
 Score: 2743.00 Matches: 529
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 10 Gaps: 0

US-09-701:395a-23 (1-529) x US-09-323-998D-22 (1-1898)

QY 1 MetGluLeuGlyValArgAsnLeuLeuSerSerCysProValTrpThrPheGlyThr 20
 Db 113 ATGGAACCTACTTGTGTTCCGAACCTCATCTCTTCTTGGCCCTGTGTGAGCTTTTGGACCA 172
 QY 21 ArgAsnLeuSerSerSerLysLeuAlaTyArgHisArgTyrGlySerSerCysArg 40
 Db 173 AGAACCTTGTAGTTTCAAAACCTAGCTTATAACATACATGATGTTCTTCTTGTAGA 232
 QY 41 ValAspPheGlnValArgAlaAspGlyGlySerGlySerArgSerValAlaTyLys 60
 Db 233 GTAGATTTTCAAGTGAGAGCTGATGTGGAGCGGGAGTAGAAGTTCTGTGTCTTATAA 292
 QY 61 GluGlyPheValAspGluGluAspPheLeuLysAlaGlyGlySerGluLeuPheVal 80
 Db 293 GAGGGTTTGTGATGAAGAGGATTTATCAAACTGGTGGTCTCTGAGCTTTTGTGTC 352
 QY 81 GlnMetGlnGlnThrLysSerMetGluLysGlnAlaLysLeuAlaAspLysLeuProPro 100
 Db 353 CAAATGCAGCAACAAGCTATGAGAGAACAGCCAGCTGCCGATGAAGTTGCCACCA 412
 QY 101 IleProPheGlyGluSerValMetAspLeuValIleGlyCysGlyProAlaGlyLeu 120
 Db 413 ATACCTTTTGGAGAACCTGTAGTGAGCTTGGTTGTAATAGGTGTGGACCTGCTGCTT 472
 QY 121 SerLeuAlaAlaGluAlaLysLeuGlyLeuLysValIleGlyCysGlyProAlaGlyLeu 140
 Db 473 TCACCTGGCTGCAGAGCTCTAGCTAGCTAGGTTGAAAGTTGGCTTATTTGGTCTGATCTT 532
 QY 141 ProPheThrAsnAsnTyrGlyValTrpGluAspGluPheLysAspLeuGlyLeuGluArg 160
 Db 533 CCTTTTACAAATAATATGCTGTGGGAAGACGAGTTCAAAGATCTTGGACCTGAAAGT 592
 QY 161 CysIleGluHisAlaTrpLysAspThrIleValTyrLeuAspAsnAlaProValLeu 180
 Db 593 TGTATCGAGCTGCTTGGAGACACCATCGTATATCTTGTAAATGATGCTCTGCTCTT 652
 QY 181 IleGlyArgAlaTyrGlyArgValSerArgHisLeuLeuHisGluGluLeuLysArg 200
 Db 653 ATTTGCTGTCATATGAGAGTATGTCGACATTTGCTACATGAGAGTTGCTGAAAGG 712
 QY 201 CysValGluSerGlyValSerTyrLeuAspSerLysValGluArgIleThrGluAlaGly 220
 Db 713 TGTGTGGAGTCAGGTGTATCATATCTTGTATCTAAAGTGAAGGATCACTGAAGCTGT 772
 QY 221 AspGlyHisSerLeuValValCysGluAsnGluIlePheIleProCysArgLeuAlaThr 240
 Db 773 GATGCCATACCTTGTAGTTGTGAAATGAGATCTTATCCCTTGCAGGCTTGCCTACT 832
 QY 241 ValAlaSerGlyAlaLaSerGlyLysLeuLeuGluTyrGluValGlyGlyProArgVal 260
 Db 833 GTTGATCTGAGCAGCTTTCAGGGAACCTTTTGGAGTATGAGTAGGTGGCCCTCGGTGT 892
 QY 261 CysValGlnThrAlaTyrGlyValGluValGluValGluAsnAsnProTyrAspProAsn 280
 Db 893 TGTGTCAAACCGCTTATGGGGTGGAGTTGAGGTGGAGAACCAATCCATCAGATCCCAAC 952
 QY 281 LeuMetValPheMetAspTyrArgAspTyrMetGlnGlnLysLeuGlnCysSerGluGlu 300
 Db 953 TTAATGCTATTCAGGACTACAGAGACTATATCAACAGAAATACAGTCTCTCGGAAGA 1012
 QY 301 GluTyrProThrPheLeuTyrValMetProMetSerProThrArgLeuPheGluGlu 320
 Db 1013 GAATATCAACATTTCTCTATGTCATGCCCATGTGCCACACAGACTTTTGTGGAGAA 1072

QY 321 ThrCysLeuAlaSerLysAspAlaMetProPheAspLeuLeuLysArgLysLeuMetSer 340
 Db 1073 ACCTGTTTGGCCCTCAAAAGATGCCATGCCATTCGATCTACTGAAGAGAAAACCTGATGCA 1132
 QY 341 ArgLeuLysThrLeuGlyIleGlnValThrLysValTyrGluGluGluTrpSerTyrIle 360
 Db 1133 CGATTGAAGACTCTGGGTATCCCAAGTTTACAAAGTTTATGAAGAGGAATGCTCATATTT 1192
 QY 361 ProValGlyGlySerLeuProAsnThrGluGlnLysAsnLeuAlaPheGlyAlaAlaAla 380
 Db 1193 CTTGTTGGTGTCTTCTTACCAACACAGAGCAAAAGAACCTAGCATTTGGTGTGAGCA 1252
 QY 381 SerMetValHisProAlaThrGlyTyrSerValValArgSerLeuSerGluAlaProLys 400
 Db 1253 AGCATGGTGCATCCCAACAGAGCTATTCGGTTGTACGGTCACTGTCAAGAGCTCCAAA 1312
 QY 401 TyrAlaSerValIleAlaLysIleLeuLysGlnAspAsnSerAlaTyrValValSerGly 420
 Db 1313 TATGCTTCTGTAATTCAAAGATTTTGAAGCAAGATTAATCTCGCTATGTGGTTCTGGA 1372
 QY 421 GlnSerSerAlaValAsnIleSerMetGlnAlaTrpSerSerLeuTrpProLysGluArg 440
 Db 1373 CAAAGTAGTGCAGTAAACATTTCAATGCAAGCATGAGCAGCTCTTTGGCCAAAGAGCGA 1432
 QY 441 LysArgGlnArgAlaPhePheLeuPheGlyLeuGluLeuIleValGlnLeuAspIleGlu 460
 Db 1433 AAACGTCAAGAGCATCTTCTTCTTTTGGATTAGAGCTTATTTGTCAGCTAGATATGAA 1492
 QY 461 AlaThrArgThrPhePheArgThrPhePheArgLeuProThrTrpMetTrpTrpGlyPhe 480
 Db 1493 GCAACGACAGAACATCTTTAGAACCTTCTTCCGCTTGGCAACTTGGATGGTGGGTTC 1552
 QY 481 LeuGlySerSerLeuSerSerPheAspLeuValLeuPheSerMetTyrMetPheValLeu 500
 Db 1553 CTTGGGTCTTCACTATCATCTTTCGATCTGCTCTTGTGTTTCCATGATGATGTTTGTG 1612
 QY 501 AlaProAsnSerMetArgMetSerLeuValArgHisLeuLeuSerAspProSerGlyAla 520
 Db 1613 GCGCCAAACAGCATGAGGATGTCATCTGTGAGACATTTGCTTTCAGATCTTCTGCTGCA 1672
 QY 521 ValMetValArgAlaTyrLeuGluArg 529
 Db 1673 GTTATGTTAGAGCTTACCTCGAAGG 1699

RESULT 2

US-09-323-998D-46
 ; Sequence 46, Application US/09323998D
 ; Patent No. US20020102631A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CUNNINGHAM JR., FRANCIS X.
 ; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND
 ; FILE REFERENCE: 108172-09019
 ; CURRENT APPLICATION NUMBER: US/09/323,998D
 ; PRIOR FILING DATE: 1999-06-02
 ; PRIOR APPLICATION NUMBER: 09/088,724
 ; PRIOR FILING DATE: 1998-06-02
 ; PRIOR APPLICATION NUMBER: 09/088,725
 ; PRIOR FILING DATE: 1998-06-02
 ; PRIOR APPLICATION NUMBER: 08/937,155
 ; PRIOR FILING DATE: 1997-09-25
 ; PRIOR APPLICATION NUMBER: 08/624,125
 ; PRIOR FILING DATE: 1996-03-29
 ; NUMBER OF SEQ ID NOS: 61
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 46
 ; LENGTH: 1848
 ; TYPE: DNA
 ; ORGANISM: Adonis palaestina
 US-09-323-998D-46

us-09-701-395a-23.rnpb

Thu May 22 09:39:17 2003

Alignment Scores:
 Pred. No.: 0 Length: 1848
 Score: 2728.00 Matches: 524
 Percent Similarity: 100.00% Conservative: 5
 Best Local Similarity: 99.05% Mismatches: 0
 Query Match: 99.45% Indels: 0
 Gaps: 10

US-09-701-395a-23 (1-529) x US-09-323-998D-46 (1-1848)

QY 1 MetGluLeuGlyValArgAsnLeuLeuSerCysProValTrpThrPheGlyThr 20
 Db 116 ATGGAACACTTGGTGTCCCACTCACTCTTCTGGCCCTGCTGGACTTTTGAACA 175
 QY 21 ArgAsnLeuSerSerLysLeuAlaTyrAsnIleHisArgTyrGlySerCysArg 40
 Db 176 AGAACCTTAGTGTCAAACTAGCTTATAACATACATCGATATGGTCTCTCTGTAGA 235
 QY 41 ValaspPheGlnValArgAlaAspGlyGlySerGlySerArgSerValAlaTyrLys 60
 Db 236 GTAGATTTCAGTGGAGGCTGATGGTGAAGCGGAGTAGAAGTCTGTGTCTTATAA 295
 QY 61 GluGlyPheValAspGluAspPheIleLysAlaGlyGlySerGluLeuLeuPheVal 80
 Db 296 GAGGTTTGTGACGAGGAGGATTTATCAAGCTGTGTGCTGTGACCTTTGTGTGTC 355
 QY 81 GlnMetGlnGlnThrLysSerMetGluLysGlnAlaLysLeuAlaAspLysLeuPro 100
 Db 356 CNAATGCGAACAAGCTATGAGAAACAGCGCAAGCTGCGGATTAAGTGTCCACCA 415
 QY 101 IleProPheGlyGluSerValMetAspLeuValIleGlyCysGlyProAlaGlyLeu 120
 Db 416 ATACCTTTCGGAATCTGTGTGAGCTTGGTGTGTAATAGTGTGGACCTGTGCTT 475
 QY 121 SerLeuAlaAlaGluAlaLysLeuGlyLeuLysValGlyLeuIleGlyProAspLeu 140
 Db 476 TCATGCTGCGAAGCTGTAGCTAGCTTGAAGTGGCTTATGCTGCTGATCTT 535
 QY 141 ProPheThrAsnAsnTyrGlyValTrpGluAspGluPheLysAspLeuGlyLeuArg 160
 Db 536 CTTTTTACAAATAATATGTGTGTGGAGAGACAGTTCAAAGATCTGACCTGAACGT 595
 QY 161 CysIleGluHisAlaTrpLysAspThrIleValTyrLeuAspAsnAspAlaProValLeu 180
 Db 596 TGTATCGAGCATGCTTGAAGGACCATCGTATATCTGACATGATGATGCTGCTT 655
 QY 181 IleGlyArgAlaTyrGlyArgValSerArgHisLeuLeuHisGluGluLeuLysArg 200
 Db 656 ATGTGCTGCTATGAGCAGTATAGCCGCAATTTGCTGATGAAGATGCTGAAAGG 715
 QY 201 CysValGluSerGlyValSerTyrLeuAspSerLysValGluArgIleThrGluAlaGly 220
 Db 716 TGTGTCGAGTCAGCTGTATCATATCTGATTTCTAAAGTGAAGGATCAGTGAAGCTGGT 775
 QY 221 AspGlyHisSerLeuValValCysGluAsnGluIlePheIleProCysArgLeuAlaThr 240
 Db 776 GATGCCATAGCTCTGTAGTTTGTGAACAGCACTTTATCCCTGCAAGCTTGTCTACT 835
 QY 241 ValAlaSerGlyAlaAlaSerGlyLysLeuGluTyrGluValGlyClyProArgVal 260
 Db 836 GTTGATCTGGAGCAGCTTCAGGAAACTTTTGGAGTATGAAGTAGGTGGCCCTCTGTGT 895
 QY 261 CysValGlnThrAlaTyrGlyValGluValGluAlaAsnAsnProTyrAspProAsn 280
 Db 896 TGTGTCCAACTGCTTATGTTGTGGAGTTGAGGTGGAGAACATCCATACGATCCCAAC 955
 QY 281 LeuMetValPheMetAspTyrArgAspTyrMetGlnLysLeuGlnCysSerGluGlu 300
 Db 956 TTAATGGTATTTATGAGCTACAGACTATATGCAACAGAAATTTACAGTGTCTGGAGAA 1015
 QY 301 GluTyrProThrPheLeuTyrValMetProMetSerProThrArgLeuPheGluGlu 320
 Db 1016 GAATATCCAACATTTCTCTATGTCATGCCATGTGCGCCACACAGACATTTTTTTTGAGAA 1075

QY 321 ThrCysLeuAlaSerLysAspAlaMetProPheAspLeuLeuLysArgLysLeuMetSer 340
 Db 1076 ACCTGTTGGCTCAAAAGATGCCATGCTTTCATCTACTGAAGAGAAAACCTAATGTCA 1135
 QY 341 ArgLeuLysThrLeuGlyIleGlnValThrLysValTyrGluGluGluTrpSerTyrIle 360
 Db 1136 CGATTGAAGACTCTGGTATCCAGTTTACAAAATTTATGAAGAGAAATGGCTTATATT 1195
 QY 361 ProValGlyGlySerLeuProAsnThrGluGlnLysAsnLeuAlaPheGlyAlaAla 380
 Db 1196 CTTGTTGGGGTCTTTACCAACACAGACAAAGAACCTAGCATTTGGTGTGTCAGCA 1255
 QY 381 SerMetValHisProAlaThrGlyTyrSerValValArgSerLeuSerGluAlaProLys 400
 Db 1256 AGCATGTGATCCAGCAACAGCTATTCGTTGTAGCATCTACTATCAGAAGCTCCAAA 1315
 QY 401 TyrAlaSerValIleAlaLysIleLeuLysGlnAspAsnSerAlaTyrValValSerGly 420
 Db 1316 TATGCTTCTGTAATTTGCAAGATTTTGAAGCAAGATACTTCGCATATCTGTTCTGGA 1375
 QY 421 GlnSerSerAlaValAsnIleSerMetGlnAlaTrpSerSerLeuTrpProLysGluArg 440
 Db 1376 CAAAGCAGTGCAGTAACATTTCAATGCAGCATGGAGCAGCTCTTGGCCAAAGAGCGCA 1435
 QY 441 LysArgGlnArgAlaPhePheLeuPheGlyLeuGluLeuIleValGlnLeuAspIleGlu 460
 Db 1436 AAACGTCAAGAGCATTTCTTCTTCTGCGGTAGAGCTTATGTGCAGCTAGATATTGAA 1495
 QY 461 AlaThrArgThrPhePheArgThrPhePheArgLeuProThrTrpMetTrpTrpGlyPhe 480
 Db 1496 GCAACAGAACAGTCTTTAGAACCTTCTCGCTGCGCAACTTGSATGTGGTGGGGTTC 1555
 QY 481 LeuGlySerSerLeuSerSerPheAspLeuValLeuPheSerMetTyrMetPheValLeu 500
 Db 1556 CTTGGGCTTCACTATCATCTTTCGATCTGATCTGTTGTTTCCATGCTACATGTTGTTT 1615
 QY 501 AlaProAsnSerMetArgMetSerLeuValArgHisLeuLeuSerAspProSerGlyAla 520
 Db 1616 GCCCGCAACAGCATGAGGATGTCATCTGTGAGACATTTGCTTTCAGATCCTTCTGTCGA 1675
 QY 521 ValMetValArgAlaTyrLeuGluArg 529
 Db 1676 GTTATGTTAAAGCTTACCTCGAAAG 1702

RESULT 3

US-09-323-998D-1
 ; Sequence 1, Application US/09323998D
 ; Patent No. US20020102631A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CUNNINGHAM JR., FRANCIS X.
 ; APPLICANT: SUN, ZAIREN
 ; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND
 ; TITLE OF INVENTION: METHODS OF USE THEREOF
 ; FILE REFERENCE: 108172-09019
 ; CURRENT APPLICATION NUMBER: US/09/323,998D
 ; PRIOR FILING DATE: 1999-06-02
 ; PRIOR APPLICATION NUMBER: 09/086,724
 ; PRIOR FILING DATE: 1998-06-02
 ; PRIOR APPLICATION NUMBER: 09/088,725
 ; PRIOR FILING DATE: 1998-06-02
 ; PRIOR APPLICATION NUMBER: 08/937,155
 ; PRIOR FILING DATE: 1997-09-25
 ; PRIOR APPLICATION NUMBER: 08/624,125
 ; PRIOR FILING DATE: 1996-03-29
 ; NUMBER OF SEQ ID NOS: 61
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 1860
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; FEATURE:
 ; NAME/KEY: CDS

LOCATION: (109)..(1680)
US-09-323-998D-1

Alignment Scores:

Pred. No.: 2,19e-228 Length: 1860
Score: 1959.50 Matches: 376
Percent Similarity: 81.65% Conservative: 60
Best Local Similarity: 70.41% Mismatches: 81
Query Match: 71.44% Indels: 17
DB: 10 Gaps: 5

US-09-701-395a-23 (1-529) x US-09-323-998D-1 (1-1860)

```

QY 1 MetGluLeuGluValArgAsnLeu-----lleSerSerCysProValTrp 16
Db 109 ATGGAGTGTGGGCTAGGAATTTCCACGATGCGGTTTCAACATTCGCTCATGG 168
QY 17 ThrPheGlyThrArgAsnLeuSerSerSerLysLeuAlaValTrpValTrp 36
Db 169 AGTTGTCGAGGAAATTTCCAGTGGTTAAGAGATACAGCTATAGGAATATTCCTTCGGT 228
QY 37 SerSerCysArgValAspPheGlnValArgAlaAspGlyCly-----SerGlySerArg 54
Db 229 ---TTGGT-----AGTGTGAGAGCTAGCGGCGGCGGAGATTTGTAAGCTT 273
QY 55 SerSerValAlaValYlyGluGlyPheValAspGluGluAspPheIleLysAlaGlyGly 74
Db 274 AGTTGTGTCGCGTGTAGAGAGATTTCCGCTGACGAGAGATTTGTAAGCTT 333
QY 75 SerGluLeuLeuPheValGlnMetGlnGlnThrLysSerMetGluLysGlnAlaLysLeu 94
Db 334 TCGAGATTCATTTGTTCAATGTCGAGACAAAGATGATGGATGATGATGATGATGAT 393
QY 95 AlaAspLysLeuProPheGlyGluSerValMetAspLeuValValIleGly 114
Db 394 GTTGATAAGTTCCTCTATATCAATGATGATGATGATGATGATGATGATGATGAT 453
QY 115 CysGlyProAlaGlyLeuSerLeuAlaGluAlaLysLeuGlyLeuLysValGly 134
Db 454 TGTGTCGTCGCTGTAGGCTTGCCTGCTGAGATTCAGCTAGCTTGTGATTAAGTTGA 513
QY 135 LeuIleGlyProAspLeuProPheThrAsnAsnTyroGlyValTrpGluAspGluPheLys 154
Db 514 CTCATTGTCGAGATTCCTCTTACTTACAAATAGCGGTTTGGGAGATGATGATCAAT 573
QY 155 AspLeuGlyLeuGluArgCysIleGluHisAlaTrpLysAspThrIleValTrpLeuAsp 174
Db 574 GATCTGGCTGCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 633
QY 175 AsnAspAlaProValLeuIleGlyArgAlaValSerArgHisLeuLeuHis 194
Db 634 GATGACAAGCTATTACCTATGCTGCTTATGGAAGATGATGATGATGATGATGAT 693
QY 195 GluGluLeuLysArgCysValGluSerGlyValSerTyroLysLeuValGlu 214
Db 694 GAGAGCTTTGAGGAGGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGT 753
QY 215 ArgIleThrGluAlaGlyAspGlyHisSerLeuValValCysGluAsnGluPheIle 234
Db 754 AGCATAACAGAGCTTCATGAGGCTTAGACTTGTCTGTGAGGACATACATACAT 813
QY 235 ProCysArgLeuAlaValAlaSerGlyAlaAlaSerGlyLysLeuLeuGluTrpGlu 254
Db 814 CCCTCAGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 873
QY 255 ValGlyGlyProArgValCysValGlnThrAlaValGlyValGluValGluAsn 274
Db 874 GTTGTGAGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 933
QY 275 AsnProTyroAspAsnLeuMetValPheMetAspTyroArgAspTyroMetGlnLys 294
Db 934 AGTCCATATGATCCAGTCAAAATGTTTTCATGGATTACAGATATATACAGAGAAA 993

```

```

QY 295 LeuGlnCysSerGluGluGluTyroProThrPheLeuTyroValMetProMetSerProThr 314
Db 994 GTTCGGAGCTTGAAGCTGAGTATCAAGCTTCTGTAGCGCATGCTATGACAAAGTCA 1053
QY 315 ArgLeuPheGluGluThrCysLeuAlaSerLysAspAlaMetProPheAspLeuLeu 334
Db 1054 AGACTCTTCTCGAGGAGCATGTTTGGCTCAAAAGATGTCATGCCCTTGTATTCCTA 1113
QY 335 LysArgLysLeuMetSerArgLysThrLeuGlyIleGlnValThrLysValTrpGlu 354
Db 1114 AAACGAAGCTCATGTTAAGATTAGATACCTCGAATTCGAATTCGAAGACTTACGAA 1173
QY 355 GluGluTrpSerTyroIleProValGlyCysLeuProAsnThrGluGlnLysAsnLeu 374
Db 1174 GAGGAGTGTCTATATCCAGTTGGTGGTTCCTTGCACAAACACGACGAAAGATCTC 1233
QY 375 AlaPheGlyAlaAlaSerMetValHisProAlaThrGlyTyroSerValValArgSer 394
Db 1234 CCTTTTGTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1293
QY 395 LeuSerGluAlaProLysTyroAlaSerValIleAlaLysLysLeuLysGlnAspSer 414
Db 1294 TTGTCTGAAGCTCAAAATATGATGATGATGATGATGATGATGATGATGATGATGAT 1353
QY 415 AlaTyroValSerGlyGlnSerSerAlaValAsnIleSerMetClnAlaTrpSerSer 434
Db 1354 AAACATCAACAGT-----ATATTTCAAGACAGCTTGGGATAT 1395
QY 435 LeuTrpProLysGluArgLysArgGlnArgAlaPhePheLeuPheGlyLeuLeuLeu 454
Db 1396 TTATGCCACCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 455
QY 455 ValGlnLeuAspIleGluAlaThrArgThrPhePheArgThrPheArgLeuProThr 474
Db 1456 GTTCAATTCGATACCAAGGATGATGATGATGATGATGATGATGATGATGATGAT 494
QY 475 TrpMetTrpTrpGlyPheLeuGlySerSerLeuSerSerPheAspLeuValLeuPheSer 515
Db 1516 TGGATGTGGCAAGGTTTCTAGGATCAACATTAACATCAGGAGATCTGCTTCTTCT 575
QY 495 MetTyroMetPheValLeuAlaProAsnSerMetArgMetSerLeuValArgHisLeuLeu 514
Db 1576 TTATCATGTTGCTCATTTCCACCAACAAATTCGAGAAAGGCTCTCATCATCTCATC 1635
QY 515 SerAspProSerGlyAlaValMetValArgAlaTyroLeuGlu 528
Db 1636 TCTGATCCACCGGACCAACCATGATAAAACCTATCTCAA 1677

```

RESULT 4

```

US-09-323-998D-24
; Sequence 24, Application US/09323998D
; Patent No. US20020102631A1
; GENERAL INFORMATION:
; APPLICANT: CUNNINGHAM JR., FRANCIS X.
; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND
; FILE REFERENCE: 108172-09019
; CURRENT APPLICATION NUMBER: US/09/323,998D
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: 09/088,724
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 09/088,725
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 08/937,155
; PRIOR FILING DATE: 1997-09-25
; PRIOR APPLICATION NUMBER: 08/624,125
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 24
; LENGTH: 1378
; TYPE: DNA

```

ORGANISM: Solanum tuberosum
 FEATURE:
 NAME/KEY: modified_base
 LOCATION: (7)..(11)
 OTHER INFORMATION: a, t, c, g, other or unknown
 FEATURE:
 NAME/KEY: modified_base
 LOCATION: (1180)..(1181)
 OTHER INFORMATION: a, t, c, g, other or unknown
 FEATURE:
 NAME/KEY: modified_base
 LOCATION: (1330)
 OTHER INFORMATION: a, t, c, g, other or unknown
 US-09-323-998D-24

Alignment Scores:
 Pred. No.: 5,61e-170 Length: 1378
 Score: 1478.50 Matches: 281
 Percent Similarity: 86.77% Conservative: 47
 Best Local Similarity: 74.34% Mismatches: 48
 Query Match: 53.90% Indels: 3
 DB: 10 Gaps: 1

US-09-701-395a-23 (1-529) x US-09-323-998D-24 (1-1378)

QY 151 AspGluPheLysAspLeuGluArgCysIleGluHisAlaTrpLys-AspThrIl 170
 Db 14 GATGAGTCAAGATCTTGGCTTCAAGCTGCATTAACATGTTGGCTGGGATACCAT 73
 QY 170 eValTyLeuAspAsnAlaProValLeuIleGlyArgAlaTyrgValSerAr 190
 Db 74 TGTATATCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 133
 QY 190 gHisLeuHisGluLeuLysArgCysValGluSerGlyValSerTyLeuAs 210
 Db 134 CCATTACTGCAGAGAGTACTCAAAAGGTGTGGAGGAGGTGTTGTATCTAAA 193
 QY 210 pSerLysValGluArgIleThrGluAlaGlyAspGlyHisSerLeuValValCysGluAs 230
 Db 194 CTCGAAAGTGGATAGGATGTTGAGGCCACAAATGGCCACAGTCTGTAGAGTGGAGGG 253
 QY 230 nGluIlePheIleProCysArgLeuAlaThrValAlaSerClyAlaSerGlyLysLe 250
 Db 254 TGATGTGTGATTCCTTCGAGGTTGTGATGATGATGATGATGATGATGATGATGATG 313
 QY 250 uLeuGluTyrgLuuValGlyGlyProArgValCysValGlnThrAlaTyrgValGluVa 270
 Db 314 CTTCAGTATGAGTGGAGGTCTAGAGTTCGTTCAACACAGCTTATGGAGTGGAGT 373
 QY 270 lGluValGluAsnAsnProTyrgAspProAsnLeuMetValPheMetAspTyrgAspTy 290
 Db 374 TGAGGTGATGAACAATCCATTGACCGGAGCCTGATGTTTCATGGATTATAGAGCTA 433
 QY 290 rMetGlnGlnLysLeuGlnCysSerGluGluTyrgProThrPheLeuTyrgValMetPr 310
 Db 434 TGTGACAGACAGCTCAATCTTTAGAGCTAAATATCCCAATTTCTCTATGCCATGCC 493
 QY 310 oMetSerProThrArgLeuPheGluThrCysLeuAlaSerLysAspAlaMetPr 330
 Db 494 CATGTCTCCAACAGAGCTTTTTCGAGGAACTTGTTCGTTCAAAAGATGCAATGCC 553
 QY 330 oPheAspLeuLeuLysArgLysLeuMetSerArgLeuLysThrLeuGlyIleGlnValTh 350
 Db 554 ATTGCGATCTCTTAAGAAAAAATGATGTTACGATTGAACACCCCTCGGTGTAAGAAATAA 613
 QY 350 rLysValTyrgLuuGluThrPserTyrgIleProValGlyGlySerLeuProAsnThrGl 370
 Db 614 AGAATTTATGAGGAGAAATGGTCTTACATACCATGTTGGAGGATCTTTGCCAAATACAGA 673
 QY 370 uGlnLysAsnLeuAlaPheGlyAlaAlaAsnSerMetValHisProAlaThrGlyTyrgSe 390
 Db 674 AAAAAAACACTTGATTTGGTCTGCTAGCATGGTTCATCCAGCCACAGGTTATTC 733

QY 390 rValValArgSerLeuSerGluAlaProLysTyrgAlaSerValIleAlaLysLeuLy 410
 Db 734 AGTCGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 793
 QY 410 sGlnAspAsnSerAlaTyrgValValSerGlyGlnSerSerAlaValAlaSerMetGl 430
 Db 794 ACATAATCATAGCAAGATATGTTACT---AGTTCAAGTACCCCGGATATTTCAACTCA 850
 QY 430 nAlaTrpSerSerLeuTrpProLysGluArgLysArgGlnArgAlaPhePheLeuG 450
 Db 851 AGCTTGGAACTCTTTGGCCACAGAACGAAACGAAAGATCGTTTCTTATTTGG 910
 QY 450 yLeuGluLeuIleValGlnLeuAspIleGluAlaThrArgThrPhePheArgThrPhePh 470
 Db 911 ACTGGCTCTGATATTCAGCTGGATATGAGGGGATAGAGTCAATTTTCGCGCGTCTT 970
 QY 470 eArgLeuProThrTrpMetTrpTrpGlyPheLeuGlySerSerLeuSerSerPheAspLe 490
 Db 971 CCGTGTGCCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1029
 QY 490 uValLeuPheSerMetTyrgMetPheValLeuAlaProAsnSerMetArgMetSerLeuVa 510
 Db 1030 CATGTTATTTGGCTTCTACATGTTTATATGACCAAAATGACATGAGAGAGGCTTAAT 1089
 QY 510 lArgHisLeuLeuSerAspProSerGlyAlaValMetValArgAlaTyrgLeu 527
 Db 1090 CAGACATCTTTTATCTGATCTCTACTGTCGCAACATGATAGAACTTATCTT 1141

RESULT 5

US-10-041-472-3
 ; Sequence 3, Application US/10041472
 ; Patent No. US20020092039A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Shemmaker, Christine
 ; TITLE OF INVENTION: METHODS FOR PRODUCING CAROTENOID COMPOUNDS AND SPECIALITY OIL
 ; FILE OF INVENTION: SEEDS
 ; FILE REFERENCE: 16516.141
 ; CURRENT APPLICATION NUMBER: US/10/041,472
 ; CURRENT FILING DATE: 2002-01-10
 ; PRIOR APPLICATION NUMBER: US 08/908758
 ; PRIOR FILING DATE: 1997-08-08
 ; PRIOR APPLICATION NUMBER: US 60/024145
 ; PRIOR FILING DATE: 1996-09-08
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 3
 ; LENGTH: 1272
 ; TYPE: DNA
 ; ORGANISM: Brassica napus
 US-10-041-472-3

Alignment Scores:
 Pred. No.: 9,8e-160 Length: 1272
 Score: 1394.00 Matches: 265
 Percent Similarity: 85.51% Conservative: 30
 Best Local Similarity: 76.81% Mismatches: 46
 Query Match: 50.82% Indels: 4
 DB: 12 Gaps: 2

US-09-701-395a-23 (1-529) x US-10-041-472-3 (1-1272)

QY 106 SerValMetAspLeu-----ValValIleGlyCysGlyProAlaGlyLeuSerLeuAla 123
 Db 67 AGTGTGATGGATATCTGCAGATTCGGCTTGTGTTGGTCTCTGCTGCTGCTGCTGCTG 126
 QY 124 AlaGluAlaAlaLysLeuGlyLeuLysValGlyLeuIleGlyProAspLeuProPheThr 143
 Db 127 GCTGAATCAGCTAAGCTTAGGACTTAAAGTTGGACTGATTGGTCTCTGACCTTCTTCACT 186
 QY 144 AsnAsnTyrgLuuValTrpGluAspGluPheLysAspLeuGlyLeuGluArgCysIleGlu 163
 Db 187 AACAACTACGGTGTGTTGGGAAAGATGATTCACACGATCTTGGCTTGCATAAATGATTGAG 246

Qy	164	HsAlaTrpLysAspThrIleValTyrLeuAsnAspAlaProValLeuIleGlyArg	183
Db	247	CATGTTTGAGAGATACCCTTGTGTATCTGGACGATCACATTCATTACCATTTGGTCGT	306
Qy	184	AlaTyrGlyArgValSerArgHisLeuLeuHisGluGluLeuLeuLysArgCysValGlu	203
Db	307	GCTTATGGAAGAGTTAGTCGACGTTACTTCCACAGAGAGTCTCTGCAGAGAGTGTGTGGAG	366
Qy	204	SerGlyValSerTyrLeuAspSerLysValGluArgIleThrGluAlaGlyAspGlyHis	223
Db	367	TCAGGTGTCGTCATCTAGCTCCCAAGTTGAGAGCATAAACAGAAGCTCCTGATGGCCTT	426
Qy	224	SerLeuValValCysGluAsnGluIlePheIleProCysArgLeuAlaThrValAlaSer	243
Db	427	AGCGTTGTTCTTCGTGCAACAACACCCTTGTCCGTGCAGGCTTGGCATGTGTCTTCT	486
Qy	244	GlyAlaAlaSerGlyLysLeuLeuGlnTyrGluValGlyGlyProArgValCysValGln	263
Db	487	GGAGCAGCTTCTGGGAAGCTCTTGCAATACGAAGTTGGAGGCCCTAGAGTCTGTGTCCA	546
Qy	264	ThrAlaTyrGlyValGluValGluValGluAsnAsnProTyrAspProAsnLeuMetVal	283
Db	547	ACTGCTTAACGGCTTGGAGGTTGAGGTGGAAAAGTAGTCCATATGATCCAGACGACATGGT	606
Qy	284	PheMetAspTyrArgAspTyrMetGlnGlnLysLeuGlnCysSerGlyGluGluTyrPro	303
Db	607	TTCATGGATTACAGAGATTATACAAACGAGAAATCCGGAGCTTAGAAGCTGATATCCA	666
Qy	304	ThrPheLeuTyrValMetProMetSerProThrArgLeuPheGluGluThrCysLeu	323
Db	667	ACGTTTTCTTACGCCATGCGCTATGACAAAGACACAGAGCTTCTTTTGAGGAGACATGTCTT	726
Qy	324	AlaSerLysAspAlaMetProPheAspLeuLeuLysArgLysLeuMetSerArgLeuLys	343
Db	727	GCCTCAAAGATGTCATGCCCTTGATTGCTTAAAAAGAGCTCTGTTGAGATTAGAG	786
Qy	344	ThrLeuGlyIleGlnValThrLysValTyrGluGluGluTyrPsrTyrIleProValGly	363
Db	787	ACACTCGGAATCCGAATACTAAAGACTTACGAAGAGGAATGCTCTTATATCCCAGTAGT	846
Qy	364	GlySerLeuProAsnThrGluGlnLysAsnLeuAlaPheGlyAlaAlaAlaSerMetVal	383
Db	847	GGTTCTCTGCCAACACGGAACAAAGAAATTCGCGCTTTGGTGTCTGAGCTAGCATGGTT	906
Qy	384	HisProAlaThrGlyTyrSerValValArgSerLeuSerGluAlaProLysTyrAlaSer	403
Db	907	CATCTCGCAACAGGCTATTACGTTGTGAGATCCTTGTCTGAAGCTCCAAAATACGCATCA	966
Qy	404	ValIleAlaLysIleLeuLysGlnAspAsnSerAlaTyrValValSerGlyGlnSerSer	423
Db	967	GTCACTGCTAAATATACATAAACAATGAGACCACTACTCTCTCTCACC-----AGACACATC	1020
Qy	424	AlaValAsnIleSerMetGlnAlaTrpSerSerLeuTyrProLysGluArgLysArgGln	443
Db	1021	AACACCAATATTTCAACACAGCTTGGGNATCTTTATGGCCACCACCAAGAAAGAACGACAG	1080
Qy	444	ArgAlaPhePheLeu	448
Db	1081	AGAGCATCTCTTTCTA	1095

RESULT 6

; US-10-041-472-2
 ; Sequence 2, Application US/10041472
 ; Patent No. US20020092039A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Shewmaker, Christine
 ; TITLE OF INVENTION: METHODS FOR PRODUCING CAROTENOID COMPOUNDS AND SPECIALITY OILS IN
 ; TITLE OF INVENTION: SEEDS
 ; FILE REFERENCE: 16516.141
 ; CURRENT APPLICATION NUMBER: US/10/041,472
 ; CURRENT FILING DATE: 2002-01-10
 ; PRIOR APPLICATION NUMBER: US 08/908758
 ; PRIOR FILING DATE: 1997-08-08

US-09-701-395A-23 (1-529) x US-09-941-947A-29 (1-1149)

QY	447	Phe---LeupPheGlyLeuGlutLeuValGlnLeuAspIleGluAlaThrArgThrPhe	465
Db	940	TTCGGATGCTGAATCGCATGTGTTTATTAGCCGCCGACCGCAGCTCACGCTGCCTGGCTGTG	999
QY	466	PheArThrPhePheAargLeuProThrTrpMetIrrPrpGlyPheLeuGlySerSerLeu	485
Db	1000	ATGACGCGTTTCTATGGCTTACCAGGATTGTGATGCCCGCTTTTATGCGGAAACTC	1059
QY	486	SerSerPheAsp 489	
Db	1060	ACCGTGACCGAT 1071	
RESULT 10			
US-09-938-842A-1578			
; Sequence 1578, Application US/09938842A			
; Patent No. US20020160378A1			
; GENERAL INFORMATION:			
; APPLICANT: Harper, Jeff			
; APPLICANT: Kieps, Joel			
; APPLICANT: Wang, Xun			
; APPLICANT: Zhu, Tong			
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS			
; TITLE OF INVENTION: SAME, AND METHODS OF USE			
; FILE REFERENCE: SCRIPT300-3			
; CURRENT APPLICATION NUMBER: US/09/938,842A			
; PRIOR FILING DATE: 2001-08-24			
; PRIOR APPLICATION NUMBER: US 60/227,866			
; PRIOR FILING DATE: 2000-08-24			
; PRIOR APPLICATION NUMBER: US 60/264,647			
; PRIOR FILING DATE: 2001-01-16			
; PRIOR APPLICATION NUMBER: US 60/300,111			
; PRIOR FILING DATE: 2001-06-22			
; NUMBER OF SEQ ID NOS: 5379			
; SEQ ID NO 1578			
; LENGTH: 1593			
; TYPE: DNA			
; ORGANISM: Arabidopsis thaliana			
US-09-938-842A-1578			
Alignment Scores:			
Pred. No.: 0.000112 Length: 1593			
Score: 123.00 Matches: 119			
Percent Similarity: 36.30% Conservative: 77			
Best Local Similarity: 22.04% Mismatches: 182			
Query Match: 4.48% Indels: 163			
DB: 9 Gaps: 27			
US-09-701-395A-23 (1-529) x US-09-938-842A-1578 (1-1593)			
QY	77	LeuLeuPheValGlnMetGlnThrLysSerMetLulysGlnAlalysLeuAlaasp	96
Db	73	CTTCTATAGCTCCTCGTCCGCGGAGCAAGCATCATGGTCTGTCAATGTCCGTAAC	132
QY	97	LysLeuProProfileproPheGlyGluSerValMetAspLeuValIleGlyCysGly	116
Db	133	GGAACCCCTAACGGTGAATCTGGAACACACGTT---GATATTATCATTTGCGTGTGCTG	189
QY	117	ProAlaglyLeuSerLeuAlaGluAlaValLysLeuGlyLeuLysValGlyLeuLle	136
Db	190	GTCCGCGCGCTGCCCTTCTCATACCCTCGCAAGGAAGAAGATTCACGTATA	249
QY	137	GlyProAspLeuProPheThrAsn---	144
Db	250	GAAGAGACTATACGAGCGCTCATCGAATVTCGGTGAATTTACTTCAGCCTGGTGTGTTAC	309
QY	145	-----AsnTyrglyValTrpGluAspGluPheLysAspLeuGlyLeuGluArg	160
Db	310	TTGAAGTTAATCGAACTCGGGCT---GAAGATTGCTGTAAGGATATAGATGCCGACAGA	366
QY	161	CystIleGluHisIala---TrpLysasp-----ThrIleValTyr---Leuasp	174
Db	367	GTCTTGTTGTTATGCTCTCTTTAAAGATGGAAACACATAACTCTCTTACCCTGGTGGAT	426

```

Db   1312 GTTTGCTCATCAGCACCTTGGCTTTGTCTATCT-GGTTTGACCACCGACTATGAGCCT 1370
Qy   490 uValLeu-----PheSerMetTyrrMepHe-                               -ValLeuAlaPr 502
      |||||    |||||     :||||          |||           |||
Db   1371 TGTTCCTCATTTCTTCGCAGTTCGAATTTCCTGGGGTTCGGTGGTTGCTTGTACTCTCC 1430
Qy   502 oAsnSerMetArgMetSerLeuValArzGHisLeuSerAspProSerGlyalaVal 521
      ||::   |||       |||             |:|||         |||
Db   1431 GTCGGTTAAACGGTTATTCGGCTTCGAGTAGCAATAAATCTCGAGTGCTTCAGGGATCATA 1488

RESULT 11
US-09-738--626-529
; Sequence 529, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: Tateishi, Naoko
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: Ozaki, Akio
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR FILING DATE: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 529
; LENGTH: 1272
; TYPE: DNA
; ORGANISM: Corynebacterium glutamicum
US-09-738-626-529

Alignment Scores:
Pred. No.:                0.000636                                Length: 1272
Score:                    115.50                                  Matches: 76
Percent Similarity:      35.69%                                   Conservative: 55
Best Local Similarity:   20.71%                                   Mismatches: 137
Query Match:              4.21%                                   Indels: 99
DB:                        9                                       Gaps: 16

US-09-701-395A-23 (1-529) x US-09-738-626-529 (1-1272)
QY   103 PheGlyGluservValMetAspLeuValVallleGlyCysglyProalaglyLeuSerLeu 122
      |||   ::   |||||     :||||          |||           |||
Db   7   TTtGTGTACAACtTTTGATgTTGAtcATCGGGCGGCCcttcAgstgcCGGCC 66
QY   123 AlaAlaGluAlaalaysLeuGlyLeu-----                          131
      |||   |||||     :||||          |||           |||
Db   67 GCcGTccATcgCGGCCAggAcTggGcTTCaaCaAttgtTatcgAGCcctTCCTttCCCG 126
QY   132 -----LysValGlyLeulle 136
      |||   CCGGATAAACGTGTGGCGATGCCTTACTCCCgTGCGATTCACcAGcTAGAACTTC TA 186
QY   137 GlyProAspLeuProPhetThrAsnAntyrGlyValTrpGluaspeluhelysAspLeu 156
      |||   |||       |||           |||           |||
Db   187 GGtGTGTGTGATcAGGTtAccGGGGATtat-----TTcaCaAGGGCTtgAAcTGcAT 240
QY   157 GlyLeuGluarGlyCystleglHisalatrPlysAspThrlleValtyrLeuAspnAsp 176
      |||   |||       |||           |||           |||
Db   241 GcTTTTGTGGcTCTGTtGAGGCGCCGTGGGAGACA-----TATTTcAGcAATAAG 294

```


Db	397		-----			---	GTGAGACCCGACCTGACG	429
QY	262	Vat	-----	GlnThrAlaTyrGlyValGluValGluValGluAsnAsnProTyrAspPro	279			
Db	430	GTGGGTTTCAGAAATTCGTGGCGTCGAGATCGAGACCGAC	---			---	GCCGCCCATGCGCTC	486
QY	280	AsnLeuMetValPheMetAsp					---TyrArgAspTyrMet	291
Db	487	GAGCGCCGATGATCATGTGACGACCGCTCCGCAGATGACGGGTACCGC				---		537
QY	292	GlnGlnLysLeuGlnCysSerGluGluGluTyrProThrPheLeuTyrValMetProMet	311			---		
Db	538	-----	TTTCATCTATCTGCTGCCCTTC	558				
QY	312	SerProThrArgLeuPhePheGluGlnThrCysLeuAlaSerLysAspAlaMetProPhe	331			---		
Db	559	AGTCCACCGCATCTGATCGAGATACGCGTACGACGACGGCGCGCATGTGACGAT	618			---		
QY	332	AspLeuLeuLysArgLysLeuMetSerArgLeuLysThrLeuGlyIleGlnValThrLys	351			---		
Db	619	GGCGCGCTGGCGCGGCTCGCTGGACTATCGCGCAGCGGGCTGGACGGCGGAG	678			---		
QY	352	ValTyrGluGluGluTrpSerTyrIleProValGly	365			---		
Db	679	ATCGCGCGGAAAGG	---	GGCATCTCTCCCATCGCGCTGGCCCATGACGCCATAGGCTTC	735			
QY	366	LeuProAsnThrGluGlnLysAsnLeuAlaPheGlyAlaAlaAlaSerMetValHisPro	395			---		
Db	736	TGGCGCGACACGCGCAGCGGGCGGTGCCGGTGGCGGACGGCTTCACGCCCC	795			---		
QY	386	AlaThrGlyTyrSerValValArgSerLeuSerGluAlaProLysTyrAlaSerValIle	405			---		
b	796	GTACCGGATATTCGCTG	-----	CCCTATGCGCGCAGGTGCGG	834			
Y	406	AlaLysIleLeuLysGlnAspAsnSerAlaTyrValSerGlyGlnSerSerAlaVal	425			---		
b	835	GATGCCATTCGGCGCGCGCACCTGACG	---	ACCGCGTCCGCG	873			
Y	426	AsnIleSerMetGlnAlaTrpSerSerLeuTrpProLysGluArgLysArgGlnArgAla	445			---		
b	874	CGTCCGCGGTGCGCGCTGGCGCATC	-----	GATCGCGGATCGGACCGCTTC	924			
Y	446	PhePheLeuPheGlyLeuGluLeuIleValGlnLeuAspIleGluAlaThrArgThrPhe	465			---		
Y	925	CTGCGGCTGTGAACCGGATGCTGTTCGCGGCTCCCGCCGACCGCTGATCGCTG	984			---		
Y	466	PheArgThrPhePheArgLeuPro	473			---		
Y	985	CTGACGCGGTTCACCGCTGCGG	1008			---		

RESULT 13

```

US-09-920-923-1/c
; Sequence 1, Application US/09920923
; Publication No. US2003002273A1
; GENERAL INFORMATION:
; APPLICANT: Pasamontes, Luis
; APPLICANT: Tsygankov, Yuri
; TITLE OF INVENTION: Improved Fermentative Carotenoid Production
; FILE REFERENCE: Improved Fermentative Carotenoid
; CURRENT APPLICATION NUMBER: US/09/920,923
; CURRENT FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 08/980,832
; PRIOR FILING DATE: 1997-12-01
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 1
; LENGTH: 8625

```

EQ ID NO 1
LENGTH: 8625

TYPE: DNA

FILE: DNA
ORGANISM: FL

ORGANISM: *Fla.*
FEATURE:

FEATURE:
NAME / KEY

NAME/KEY: uns

LOCATION: (83)

1

—

```

; NAME/KEY: unsure
; LOCATION: (8539) .. (8540)
; NAME/KEY: unsure
; LOCATION: (8581)
; NAME/KEY: unsure
; LOCATION: (8590)
; NAME/KEY: unsure
; LOCATION: (8592)
; NAME/KEY: unsure
; LOCATION: (8602) .. (8604)
US-09-920-923-1

```

Alignment Scores:

Pred. No.:	0.0284	Length:	8625
Score:	113.00	Matches:	83
Percent Similarity:	33.82%	Conservative:	55
Best Local Similarity:	20.34%	Mismatches:	152
Query Match:	4.12%	Indels:	118
DB:	9	Gaps:	18

US-09-701-395A-23 (1-529) x US-09-920-923-1 (1-8625)

QY	109	AspLeuValValIleGlyCysGlyProAlaGly	-----LeuSerLeuAlaAlaGluAla	126
Db	6933	GATCTGTTGATCGCGGCGGGGCTGTCGGT	CGCTGATCGCGCTGCGCTTCGCTCGCAC	6876
QY	127	AlaLysLeuGlyLeuLysValGlyLeuLeu	-----GlyPro	138
		:::		
Db	6873	CGACACCGGATCGCGGCATCGTGTGATGCT	GACGCGGGTCCGGCCCTCGGCACGACGAC	6814
QY	139	-----AspLeuProPheThrAsnAsnTyr	-----	146
Db	6813	ACCTGGTCTTCGCACGACGAGATCTTTCGCC	CGGATGGTGGCGCGCGCTGCGCCCAT	6754
QY	147	-----GlyValTrpGluAspGlu	-----PheLysAspLeuGlyLeuGluArgCys	161
Db	6753	CGTCGCGCGAATGAGCAGCGATCAGAGGCTG	CGCGTTCCTCCGAC	6712
QY	162	IleGluHisAlaTrpLysAspThrIleValTyrLeu	AspAsnAspAlaProValLeuIle	181
Db	6711	-----CATTCGCGCGGCTGACGACAGGCTATG	CGCTGATCGATCGAGCGGGCGCGCTGATC	6658
QY	182	GlyArgAlaTyrGlyArgValSerArgHisLeuLeu	HisGluGluLeuLysArgCys	201
Db	6657	GGGCTCTCGAGGTGTGATCTCGCGTGGATACG	CACTAT	6619
QY	202	ValGluSerGlyValSerTyrLeuAspSerLysVal	GluArgIleThrGluAlaGlyAsp	221
b	6618	-----GTTCGGCGGCTGGAGGATACCGCGCGG	ACGCTGACGGACGGCTCGCGG	6571

QY	222	GlyHisSerLeuValValCysGluAsnGluIlePheIleProCysArgLeuAlaThrVal	241
Db	570	ATCGAGGCTGCCTGGCTGATCGAC	
QY	242	AlaSerGlyAlaAlaSerGlyLysLeuLeuGluTyrGluValGlyGlyProArgValCys	261
Db	646	GCCCGTGGTGC-----GTCGAGACCCCGCACCTGGACC	
QY	262	Val-----GlnThrAlaTyrGlyValGluValGluValGluAsnAsnProTyrAspPro	279
Db	6513	GTGGTTCACAGAAATTCGTGGCGTCGATCGAGACCGAC-----GCCCCCATGGCGTC	6457
QY	280	AsnLeuMetValPheMetAsp-----TyrArgAspTyrMet	291
Db	6456	GAGCGCCCGATGATCATGGACGCGACCGTTCGCGAGATGACGGGTACCGC-----	6406
QY	292	GlnGlnLysLeuGlnCysSerGlnGluGluTyrProThrPheLeuTyrValMetProMet	311
Db	6405	-----TTTCATCTATCTGCTGCCCTTC	6385
QY	312	SerProThrArgLeuPhePheGluGluThrCysLeuAlaSerLysAspAlaMetProphe	331
Db	6384	AGTCCACCCGCATCTCTGATCGAGGATACCGCTACAGCGAGCGCGCGCATCTGGACGAT	6325

us-09-701-395a-23.rnpb

Thu May 22 09:39:17 2003

QY 139 -----AspLeuProPheThrAsnAsnTyr----- 146
 Db 7495 ACCTGCTCTGCCAGCACCGATCTTTCGCCGAATGCTGGCGCCCTGTCGCCCAT 7436
 QY 147 -----GlyValTrpGluAspGlu-----PheLysAspLeuGlyLeuGluArgCys 161
 Db 7435 CGTCGCGCGGAATGACGAGTATGCTGCGGCTTTCCTGAC----- 7394
 QY 162 IleGluHisAlaTrpLysAspThrIleValTyrLeuAspAsnAlaProValLeuIle 181
 Db 7393 -----CATTCGCGCGCTGACGACGAGTATGCTGACGAGCGCGGCGCTGATC 7340
 QY 182 GlyArgAlaTyrGlyArgValSerArgHisLeuLeuHisGluGluLeuLysArgCys 201
 Db 7339 GGGCTGTCGAGGCTGTCGATCTGCGGTGATACGAT----- 7301
 QY 202 ValGluSerGlyValSerTyrLeuAspSerLysValGluArgIleThrGluAlaGlyAsp 221
 Db 7300 -----GTCGCGAGCTGGACGATACCGCGCGACGCTGACGAGCGCTCGCG 7253
 QY 222 GlyHisSerLeuValCysGluAsnGluIlePheIleProCysArgLeuAlaThrVal 241
 Db 7252 ATCGAGGCTGCTCGCTGCGTATCGAC-----GTCGAGACCGCGACCTGACC 7196
 QY 242 AlaSerGlyAlaAlaSerGlyLysLeuLeuGluTyrGluValGlyGlyProArgValCys 261
 Db 7228 GCCGTGGTGCC-----GlnThrAlaTyrGlyValGluValGluAsnAsnProTyrAspPro 279
 QY 262 Val-----GlnThrAlaTyrGlyValGluValGluValGluAsnAsnProTyrAspPro 279
 Db 7195 GTGGGTTTCAGAAATTCGTGGCGCTCGAGATCGAGACCGAC-----GCCCGCATGCGGCTC 7139
 QY 280 AsnLeuMetValPheMetAsp-----TyrArgAspTyrMet 291
 Db 7138 GAGCGCGCGATGATATGACGCGACCGCTCCGAGATGAGCGGTACCGC----- 7088
 QY 292 GlnGlnLysLeuGlnCysSerGluGluGluTyrProThrPheLeuTyrValMetProMet 311
 Db 7087 -----TTCATCTATCTGCTGCCCTTC 7067
 QY 312 SerProThrArgLeuPheGluGluThrCysLeuAlaSerLysAspAlaMetProPhe 331
 Db 7066 AGTCCGACCGCATCTGATCGAGGATACGCTACGACGCGCGGCGATCTGGACGAT 7007
 QY 332 AspLeuLysArgLysLeuMetSerArgLysThrLeuGlyIleGlnValThrLys 351
 Db 7006 GCGCGCTGCGCGAGGCGTCCCTGGACTATGCCCGACGCGGCTGGACCGGCGAGGAG 6947
 QY 352 ValTyrGluGluGluTyrSerTyrIleProValGly-----GlySer 365
 Db 6946 ATGCGCGCGAAAGG---GGCATCTCTGCCCATCGCTGCCCATGACGCGATAGGCTTC 6890
 QY 366 LeuProAsnThrGluGlnLysAsnLeuAlaPheGlyAlaAlaAlaSerMetValHisPro 385
 Db 6889 TGGCGCGACCGCGCGGCGGCGTCCGCTGGGCTGGGCGGCGGCGCTTCCACCCC 6830
 QY 386 AlaThrGlyTyrSerValValArgSerLeuSerGluAlaProLysTyrAlaSerValIle 405
 Db 6829 GTCACCGGATATTCGCTG-----CCCATGCGCGCGAGTCCGCG 6791
 QY 406 AlaLysIleLeuLysGlnAspAsnSerAlaTyrValSerGlyGlnSerAlaVal 425
 Db 6790 GATGCCATCCGCGCGCGGCGACCTGACG-----ACCGCGTCCGCC 6752
 QY 426 AsnIleSerMetGlnAlaTrpSerSerLeuTyrProLysGluArgLysArgGlnAla 445
 Db 6751 CGTCGCGGCTGCGCGCTGGCGCATC-----GATCGCGGATCGCGACGCGCTC 6701
 QY 446 PhePheLeuPheGlyLeuGluLeuIleValGlnLeuAspIleGluAlaThrArgPhe 465
 Db 6700 CTGCGGCTGCTGAACCGGATGCTGTTCCGCGGCTGCCCGCGCGACCGCTGCTATCGCTG 6641

QY 332 AspLeuLysArgLysLeuMetSerArgLeuLysThrLeuGlyIleGlnValThrLys 351
 Db 6324 GCGCGCTGCGCGAGCGCTGCTGACTATCGCGCAGCGGCGCTGACCGCGGCGAGGAG 6265
 QY 352 ValTyrGluGluTyrSerTyrIleProValGly-----GlySer 365
 Db 6264 ATCGCGCGGAAAGG---GGCATCTCTGCCCATCGCTGCGGCGGCGGCGGCTGCTC 6208
 QY 366 LeuProAsnThrGluGlnLysAsnLeuAlaPheGlyAlaAlaAlaSerMetValHisPro 385
 Db 6207 TGGCGCGACCGCGCGGCGGCGTGGCTGGGCGGCGGCGGCTGTCACCGCC 6148
 QY 386 AlaThrGlyTyrSerValValArgSerLeuSerGluAlaProLysTyrAlaSerValIle 405
 Db 6147 GTCACCGGATATTCGCTG-----CCCATGCGCGCGAGTCCGCG 6109
 QY 406 AlaLysIleLeuLysGlnAspAsnSerAlaTyrValSerGlyGlnSerAlaVal 425
 Db 6108 GATGCCATCGCGCGCGCGCGCTGACG-----ACCGCGTCCGCC 6070
 QY 426 AsnIleSerMetGlnAlaTrpSerSerLeuTyrProLysGluArgLysArgGlnAla 445
 Db 6069 CCGCGCGGCTGCGCGCTGCGGCGATC-----GATCGCGGCGATCGCGCGCTTC 6019
 QY 446 PhePheLeuPheGlyLeuGluLeuIleValGlnLeuAspIleGluAlaThrArgPhe 465
 Db 6018 CTGCGGCTGCTGACCGGATGCTGTCGCGGCTGCCCGCGCGCGCTGCTGCGCTG 5959
 QY 466 PheArgThrPhePheArgLeuPro 473
 Db 5958 CTGACGCGGTCTACGCGCTGCCG 5935

RESULT 14
 US-09-920-923-27/c
 ; Sequence 27, Application US/09920923
 ; Publication No. US2003002273A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Pasamontes, Luis
 ; TITLE OF INVENTION: Improved Fermentative Carotenoid Production
 ; FILE REFERENCE: Improved Fermentative Carotenoid
 ; CURRENT APPLICATION NUMBER: US/09/920,923
 ; PRIORITY FILING DATE: 2001-08-02
 ; PRIOR APPLICATION NUMBER: 08/980,832
 ; NUMBER OF SEQ ID NOS: 66
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 27
 ; LENGTH: 11233
 ; TYPE: DNA
 ; ORGANISM: Unknown
 ; FEATURE:
 ; OTHER INFORMATION: Description of Unknown Organism: Plasmid pZea4
 US-09-920-923-27

Alignment Scores:
 Pred. No.: 0.0435 Length: 11233
 Score: 113.00 Matches: 83
 Percent Similarity: 33.82% Conservations: 55
 Best Local Similarity: 20.34% Mismatches: 152
 Query Match: 4.12% Indels: 118
 DB: 9 Gaps: 18

US-09-701-395A-23 (1-529) x US-09-920-923-27 (1-11233)
 QY 109 AspLeuValIleGlyCysGlyProAlaGly-----LeuSerLeuAlaGluAla 126
 Db 7615 GATCTGTCGTCGCGGCGGCGGCTGCTGCGGTCGCTGATCGGCTGCGGTCGCGAC 7556
 QY 127 AlaLysLeuGlyLeuLysValGlyLeuIle-----GlyPro----- 138
 Db 7555 GCGACGCGGATGCGCGATGCTGATGCTGACGCGCGGCTGCCCGCTCGGACGACGAC 7496

```

QY 466 PheArgThrPhePheArgLeuPro 473
Db 6640 CTGACGCGTCTTACCGCTGCCG 6617
RESULT 15
US-09-939-964-1/c
; Sequence 1, Application US/09939964
; Publication No. US20030054522A1
; GENERAL INFORMATION:
; APPLICANT: Rosenthal, Andre
; APPLICANT: Freiberg, Christoph
; APPLICANT: Perret, Xavier Philippe
; APPLICANT: Broughton, William John
; TITLE OF INVENTION: Genomic Sequence of Rhizobium SP. NGR 234 Symbiotic
; FILE REFERENCE: CARP0068
; CURRENT APPLICATION NUMBER: US/09/939,964
; PRIOR FILING DATE: 2001-08-27
; PRIOR FILING DATE: 1999-06-22
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 1
; LENGTH: 536165
; TYPE: DNA
; ORGANISM: Rhizobium
US-09-939-964-1

Alignment Scores:
Pred. No.: 34.5 Length: 536165
Score: 111.50 Matches: 53
Percent Similarity: 41.28% Conservative: 44
Best Local Similarity: 22.55% Mismatches: 106
Query Match: 4.06% Indels: 32
DB: 9 Gaps: 9

US-09-701-395a-23 (1-529) x US-09-939-964-1 (1-536165)
QY 54 ArgSerSerValAlaTyrLysGluGly-PheValAspGluGluAspHe-----IleY 71
Db 295477 CGACGATCTATCGATTACCGCGGTCTTTCAGGAGGACTTCTACGCGCCCTGATAA 295418
QY 71 salaclyGlySerGluLeuLeuPheValGlnMetGlnGlnThrLysSerMetGluLysG 91
Db 295417 CAAGGCAACGGTGGTCTTCTCTCGTACGCGCGCTGCTGCGCGCGCGCGGACA 295358
QY 91 nAlaLysLeu-----AlaAspLysLeuProIleProPheG 104
Db 295357 GGTTCCTTGGTCGCGGATACCGTCTGTCGAGCGATATGCTCTCGCGACTGCCTGTCG 295298
QY 104 yGluSerValMetAsp-----LeuValIleGlyCysGlyProAlaGlyLeuSerLe 122
Db 295297 GAGTGGTGTGAAAGCAAAAGGTTGTTCTCGTAGGGTGGCGGCCATCGCGACTTTCGC 295238
QY 122 uAlaAlaGluAlaAlaLysLeuGlyLeu---LysValGlyLeuIleGlyProAspLeuPr 141
Db 295237 TCGAGTCGAGCTGTCGCGCTGAGCGCTGAGCGCTGAGCGCTGAGCGCTGAGCGCTG 295178
QY 141 oPheThrAsnAsnTyrGlyValTyrGluAspGluPheLysAspLeuGlyLeuGluArgCy 161
Db 295177 CGAGCCTGGCAACACGGTACATGG-----CGCGTGGTGGTTC 295139
QY 161 stleGluHisAlaTrpLysAspThrIleVal-----TyrLeuAspAsnAlaPro-- 178
Db 295138 GGTGGGGGCGCTGCAAAAGACAAACCGCGCTGAGCGCTGAGCGCTGAGCGCTGAGCG 295079
QY 179 -ValLeuIleGlyArgAlaGlyArgValSerArgHisLeuLeuHisGluGluLeuLe 198
Db 295078 GACAAATGTGCGGCGTGGACCGCTAAAGTGGGCTCAGCAATCTCGAATGTGGACGAGT 295019
QY 198 uLysArgCysValGluSerGlyValSerTyrLeuAspSerLysValGluArgIleThrGl 218
Db 295018 CCGCAACTCGAGGGAATCCCATCGTGGCTTCGAGCATTTGATCGAGGACCGCGACAT 294959

```

```

QY 218 uAlaGlyAspGlyHisSerLeuValValCysGluAsnGluIlePheIleProCysArg-- 237
Db 294958 TGTGGTAGACACTTCGGCTCGACCGAGTGTACAGGCGCCCTTGGCCTACATGTCCGTAG 294899
QY 238 -----LeuAlaThrValAlaSerGlyAlaAlaSerGlyAlaAlaSerGlyLysLeuLe 251
Db 294898 CATAGGAAAGCGGTACTCTACGCGCCACGCAACGAGCGCGCGCGGAGCGCTTGTGGC 294839
QY 251 uGluTyrGluValGlyGlyPro-----ArgValCysValGln 263
Db 294838 GAGATTCAAACCGGAGCGCGGCTTGTATGCTGCTGTGCAG 294796

```

Search completed: May 22, 2003, 01:22:38
Job time : 573 secs

us-09-701-395a-23.rni

Thu May 22 09:39:17 2003

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 21, 2003, 22:18:48 ; Search time 68 Seconds
(without alignments)
2385.765 Million cell updates/sec

Title: us-09-701-395A-23
Perfect score: 2743
Sequence: 1 MELLGVRNLSSCPVWTFCT.....VRHLLSDPSGAVWVAYLER 529

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=framet_p2n.model
-Q=/cgn2_1/USPTO_spool/US09701395/runat_15052003.165752.3396/app_query.fasta_1.711
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCT=0
-LOOPEXT=0 -UNITS=bits -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09701395 -CGN1_1_40 -runat_15052003.165752.3396 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXI=7

Database :

Issued Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	1959.5	71.4	1860	1	US-08-624-125-1 Sequence 1, Appl
2	1934	70.5	1887	4	US-09-201-641-5 Sequence 5, Appl
3	1394	50.8	1272	4	US-08-908-758-3 Sequence 3, Appl
4	1202	43.8	960	4	US-08-908-758-2 Sequence 2, Appl
5	750.5	27.4	1650	1	US-08-399-561-5 Sequence 5, Appl
6	750	27.3	1959	4	US-09-201-641-1 Sequence 1, Appl
7	739	26.9	1590	4	US-08-908-758-4 Sequence 4, Appl
8	737	26.9	1614	1	US-08-399-561-4 Sequence 4, Appl
9	722.5	26.3	2897	4	US-09-134-607A-11 Sequence 11, Appl
10	722.5	26.3	3265	4	US-09-134-607A-15 Sequence 15, Appl
11	719.5	26.2	1666	4	US-09-134-607A-8 Sequence 8, Appl
12	719.5	26.2	1666	4	US-09-134-607A-13 Sequence 13, Appl

13	719.5	26.2	2876	4	US-09-134-607A-9 Sequence 9, Appl
14	719.5	26.2	2876	4	US-09-134-607A-14 Sequence 14, Appl
15	705.5	25.7	1756	2	US-08-702-598-1 Sequence 1, Appl
16	703.5	25.6	1740	4	US-09-134-607A-10 Sequence 10, Appl
17	703.5	25.6	1740	4	US-09-134-607A-12 Sequence 12, Appl
18	703.5	25.6	1740	4	US-09-134-607A-16 Sequence 16, Appl
19	606	22.1	4928	1	US-08-399-561-1 Sequence 1, Appl
20	137.5	5.0	1235	1	US-08-095-726-13 Sequence 13, Appl
21	137.5	5.0	1235	1	US-08-096-623A-13 Sequence 13, Appl
22	137.5	5.0	1893	3	US-09-147-009-5 Sequence 5, Appl
23	134.5	4.9	1235	1	US-08-095-726-15 Sequence 15, Appl
24	134.5	4.9	1235	1	US-08-096-623A-15 Sequence 5, Appl
25	129	4.7	1161	1	US-08-663-310-5 Sequence 5, Appl
26	129	4.7	1161	2	US-09-006-491-5 Sequence 5, Appl
27	129	4.7	1161	3	US-09-335-919-5 Sequence 7, Appl
28	129	4.7	2886	1	US-08-663-310-7 Sequence 7, Appl
29	129	4.7	2886	2	US-09-006-491-7 Sequence 7, Appl
30	129	4.7	2886	3	US-09-335-919-7 Sequence 9, Appl
31	125.5	4.6	1149	1	US-07-783-705A-9 Sequence 13, Appl
32	125.5	4.6	6918	1	US-07-783-705A-13 Sequence 8, Appl
33	113	4.1	1149	3	US-08-660-645A-8 Sequence 8, Appl
34	113	4.1	1149	3	US-09-298-718-6 Sequence 8, Appl
35	113	4.1	1149	4	US-09-546-969-8 Sequence 1, Appl
36	113	4.1	8625	4	US-08-980-832-1 Sequence 3, Appl
37	113	4.1	1233	4	US-08-980-832-27 Sequence 27, Appl
38	112	4.1	1748	3	US-09-147-009-3 Sequence 3, Appl
39	109.5	4.0	1650	2	US-08-500-635A-11 Sequence 11, Appl
40	109.5	4.0	1650	4	US-09-167-151-11 Sequence 11, Appl
41	109.5	4.0	4656	4	US-09-425-665-1 Sequence 1, Appl
42	109.5	4.0	4656	4	US-09-685-668-1 Sequence 1, Appl
43	104.5	3.8	1756	3	US-09-147-009-1 Sequence 1, Appl
44	99.5	3.6	2443	2	US-08-745-934-2 Sequence 2, Appl
45	99.5	3.6	11340	4	US-08-961-527-147 Sequence 147, App

ALIGNMENTS

RESULT 1
US-08-624-125-1
; Sequence 1, Application US/08624125

; Patent No. 574341
; GENERAL INFORMATION:

; APPLICANT: CUNNINGHAM JR., FRANCIS X.

; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND

; TITLE OF INVENTION: METABOLISM AND A SYSTEM FOR SCREENING SUCH GENES

; NUMBER OF SEQUENCES: 21

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,

; CITY: ARLINGTON

; STATE: VA

; COUNTRY: USA

; ZIP: 22202

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/624,125

; FILING DATE: 29-MAR-1996

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: KELBER, STEVEN B.

; REGISTRATION NUMBER: 30,073

; REFERENCE/DOCKET NUMBER: 2747-063-27

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 703-413-3000

; TELEFAX: 703-413-2220

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1860 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 109..1680
 ; OTHER INFORMATION: /product= "E-CYCLASE FROM A.
 ; OTHER INFORMATION: THALIANA"
 US-08-624-125-1

Alignment Scores:

Pred. No.: 3,77e-220 Length: 1860
 Score: 1959.50 Matches: 376
 Percent Similarity: 81.65% Conservative: 60
 Best Local Similarity: 70.41% Mismatches: 81
 Query Match: 71.44% Indels: 17
 DB: 1 Gaps: 5

US-09-701-395a-23 (1-529) x US-08-624-125-1 (1-1860)

QY 1 MetGluLeuLeuGlyValArgAsnLeu-----IleSerSerCysProValTrp 16
 DB 109 ATGGAGTGTGGGGCTAGGAATTCGAGCAATGGCGGTTTCAACATTTCCGTCATGG 168
 QY 17 ThrPheGlyThrArgAsnLeuSerSerSerLeuAlaTyrAsnIleHisArgTyrGly 36
 DB 169 ACTTGTGGAAGAAATTTCCAGTGTGTTAAGACATACAGCTATAGGAATATTCGTTTCGGT 228
 QY 37 SerSerCysArgValAspPheGlnValArgAlaAspGlyGly-----SerGlySerArg 54
 DB 229 -----TTGTGT-----AGTGTGAGCTAGCGCGGGGGAAGTTCGCTAGTGAG 273
 QY 55 SerSerValAlaTyrLysGluGlyPheValAspGluGluAspPheIleLysAlaGlyGly 74
 DB 274 AGTTGTGAGCGGTGAGAGAGATTTGCTGACGAGAGAGATTTGTGAAGCTGGTGT 333
 QY 75 SerGluLeuLeuPheValGlnMetGlnGlnThrLysSerMetGluLysGlnAlaLysLeu 94
 DB 334 TCTGAGATTTCTATTGTTCAATGTCAGCAGCAACAAAGATATGGATGAACAGTCTAAGCTT 393
 QY 95 AlaAspLysLeuProIleProIlePheGlyGluSerValMetAspLeuValIleGly 114
 DB 394 GTTGATAGTTCGCCCTATATCAATGTTGATGTTGCTGATGCTGATGCTGATGCTGATGCT 453
 QY 115 CysGlyProAlaGlyLeuSerLeuAlaGluAlaLysLeuGlyLeuLysValGly 134
 DB 454 TGTGCTCTGCTGGTTAGCTTGGCTGCGAGATCAGTAAGCTTGGATTAAGATTTGA 513
 QY 135 LeuIleGlyProAspLeuProPheThrAsnAsnTyrGlyValTrpGluAspGluPheLys 154
 DB 514 CTCATTGTCAGATCTCTCTTTACTAACAAATACGGTGTGTTGGGAAGATGAATCAAT 573
 QY 155 AspLeuGlyLeuGluArgCysIleGluHisAlaTyrLysAspThrIleValTyrLeuAsp 174
 DB 574 GATCTGGGCTGCAAAATGATTGAGCATGTTTCGAGAGAGACTATTGCTGATCTGAT 633
 QY 175 AsnAspAlaProValLeuIleGlyArgAlaTyrGlyArgValSerArgHisLeuHis 194
 DB 634 GATGACAGCTATTACCATGGCGCTGCTTATGGAAGATGATGAGCTTGTGCTCAT 693
 QY 195 GluGluLeuLeuLysArgCysValGluSerGlyValSerTyrLeuAspSerLysValGlu 214
 DB 694 GAGGAGCTTTTGGAGGCTGCTGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 753
 QY 215 ArgIleThrGluAlaGlyAspGlyHisSerLeuValValCysGluAsnGluIlePheLe 234
 DB 754 AGCATTAACAGAGCTCTGATGGCTTATGAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCT 813
 QY 235 ProCysArgLeuAlaThrValAlaSerGlyAlaAlaSerGlyLysLeuLeuGluTyrGlu 254
 DB 814 CCTGTCAGGCTGCGCAGTGTCTCTGAGCAGCTTCGGGAAGCTCTTGAATACGAA 873

QY 255 ValGlyGlyProArgValCysValGlnThrAlaTyrGlyValGluValGluAsn 274
 DB 874 GTTGTGGACCTAGAGTCTGTGTGCAAACTGCATACGGCTGAGGTTGAGGTGGAAAT 933
 QY 275 AsnProTyrAspProAsnLeuMetValPheMetAspTyrArgAspTyrMetGlnLys 294
 DB 934 AGTCCATATGATCCAGATCAAAATGTTTTCATGGATTACAGAGATTATATAACGAGAA 993
 QY 295 LeuGlnCysSerGluGluGluTyrProThrPheLeuTyrValMetProMetSerProThr 314
 DB 994 GTTCGGAGCTTAGAAGCTGAGTATCCAGCTTCTGTACGCCATGCTATGACAAGTCA 1053
 QY 315 ArgLeuPheGluGluThrCysLeuAlaSerLysAspAlaMetProPheAspLeuLeu 334
 DB 1054 AGACCTCTTCTCGAGAGACATGTTGGCTCAAAAGATGTCATGCCCTTTCATTTGCTA 1113
 QY 335 LysArgLysLeuMetSerArgLeuLysThrLeuGlyIleGlnValThrLysValTyrGlu 354
 DB 1114 AAACGAAAGCTCATGTTAAGATTAGATACACTCGGAATTCGAATCTAAAGACTTACGAA 1173
 QY 355 GluGluTrpSerTyrIleProValGlySerLeuProAsnThrGluGlnLysAsnLeu 374
 DB 1174 GAGGAGTGTCTATATCCAGTGTGTTGCTTCCCAACACCGCAACAAAGATCTC 1233
 QY 375 AlaPheGlyAlaAlaSerMetValHisProAlaThrGlyTyrSerValValArgSer 394
 DB 1234 GCCTTTGTTGCTGCGCTAGCATGTCATCCGCAACAGGCTATTCAGTTGTGAGATCT 1293
 QY 395 LeuSerGluAlaProLysTyrAlaSerValIleAlaLysIleLeuLysGlnAspLysSer 414
 DB 1294 TTGCTGAAGCTCCAAATATATCATCAGTCATCGCAGAGATACATAAGAGAGAGACTACC 1353
 QY 415 AlaTyrValValSerGlyGlnSerSerAlaValAsnIleSerMetGlnAlaTrpSerSer 434
 DB 1354 AAACAGATCAACAGT-----ATATTTCAAGACAGCTTGGGATCT 1395
 QY 435 LeuTrpProLysGluArgLysArgGlnArgAlaPhePheLeuPheGlyLeuGluLeu 454
 DB 1396 TTATGCCACCAAGAAAGAAAGACAGAGCATTTCTTCTCTTGTGCTTGCACCTCAT 1455
 QY 455 ValGlnLeuAspIleGluAlaThrArgThrPhePheArgThrPhePheArgLeuProThr 474
 DB 1456 GTTCAATTCGATCGAAGGCAATAGAACCTTCTTCCGCTTCTTCCGCTTCCAAA 1515
 QY 475 TrpMetTrpTrpGlyPheLeuGlySerSerLeuSerSerPheAspLeuValLeuPheSer 494
 DB 1516 TGGATGTGCAAGGTTCTTAGGATCAACATTAACATCAGGAGATCTGCTTCTTGTCT 1575
 QY 495 MetTyrMetPheValLeuAlaProAsnSerMetArgMetSerLeuValArgHisLeuLeu 514
 DB 1576 TTATACATGTTGCTCATTTCCAAACAATTTGAGAAAGGTCATCATCATCATCTCATC 1635
 QY 515 SerAspProSerGlyAlaValMetValArgAlaTyrLeuGlu 528
 DB 1636 TCTGATCAACCGGAGCAACCATGATAAAACCTATCTCAAA 1677

RESULT 2

US-09-201-641-5
 ; Sequence 5, Application US/09201641A
 ; Patent No. 6232530
 ; GENERAL INFORMATION:
 ; APPLICANT: Cunningham Jr, Francis X
 ; APPLICANT: DellaPenna, Dean
 ; TITLE OF INVENTION: Method for Regulating Carotenoid Biosynthesis in
 ; TITLE OF INVENTION: Mariogolds
 ; FILE REFERENCE: Quest 41-162
 ; CURRENT APPLICATION NUMBER: US/09/201,641A
 ; CURRENT FILING DATE: 1998-11-30
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 5
 ; LENGTH: 1887

; TYPE: DNA
 ; ORGANISM: Tagetes erecta
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (141)..(1688)
 ; OTHER INFORMATION: epsilon-cyclase
 US-09-201-641-5

Alignment Scores:

Pred. No.: 3,85e-217 Length: 1887
 Score: 1934.00 Matches: 365
 Percent Similarity: 83.95% Conservativeness: 64
 Best Local Similarity: 71.43% Mismatches: 74
 Query Match: 70.51% Indels: 8
 DB: 4 Gaps: 2

US-09-701-395a-23 (1-529) x US-09-201-641-5 (1-1887)

QY 22 AsnLeuSerSerSerLysLeuAlaTyrAsnIleHisArgTyrGlySerSerCysArgVal 41
 Db 159 CACATGACGGCAACAATGGCGCTTTTACATGCCCTAGGTTTATCAGTATCAGATCAGATAC 218
 QY 42 AspPheGlnValArgAlaAspGlySerGlySerArgSerSerValAlaTyrLysGlu 61
 Db 219 ACGAAGCAATTAAGTGC-----AACGCTCTAAAGCCAGCTAGTCTGTTAAACAA 269
 QY 62 GlyPheValAspGluGluAspPheIleLysAlaGlyGlySerGluLeuLeuPheValGln 81
 Db 270 GAGATTGAGGAGGAAGATTATGTCAAGCCGGTGGATCGGAGCTGCTTTTGTTCAA 329
 QY 82 MetGlnGlnThrLysSerMetGluGlnAlaLysLeuAlaAspLysLeuProPhe 101
 Db 330 ATGCAACAGAAATAGTCCATGGATGCAGCTAGCTAGCTATCCCAAAAGCTCCCAAGGGA 389
 QY 102 PropheGly-----GluSerValMetAspLeuValIleGlyCysGly 116
 Db 390 CCATAGAGGAGGAGGAGACAGTACTGATCTGATGTTGGTTGTAATTTGGTTGGT 449
 QY 117 ProAlaGlyLeuSerLeuAlaAlaGluAlaLysLeuGlyLeuLysValGlyLeuIle 136
 Db 450 CCTGCTGGCTTGTCTTGTCTGGAGAAATCAGCAAGCTAGGCTTGAATTCGCCACTATC 509
 QY 137 GlyProAspLeuProPheThrAsnAsnTyrGlyValrPGLuAspGluPheLysAspLeu 156
 Db 510 GGCCCTCATCTCTCTTTTCAAAATTAAGTATGTTGGGAGGATGAATTTATAGTCTT 569
 QY 157 GlyLeuGluArgCysIleGluHisAlaTrpLysAspThrIleValTyrLeuAspAsnAsp 176
 Db 570 GGACTTGAGGCTGATTGACATGTTGGCGAGATCTGTAGTATATCTTGATGACAAAC 629
 QY 177 AlaProValLeuIleGlyArgAlaTyrGlyArgValSerArgHisLeuLeuHisGluGlu 196
 Db 630 GATCCCATCTCATAGGTCGCTGATGGAGGAGTACTGCTGATTTACTTACGAGGAG 689
 QY 197 LeuLeuLysArgCysValGluSerGlyValSerTyrLeuAspSerLysValGluArgIle 216
 Db 690 TTGTTGACTAGGTGATGAGTCAAGCGTTTCATATCTGAGCTCCAAAGTGAAGCGGATT 749
 QY 217 ThrGluAlaGlyAspGlyHisSerLeuValValCysGluAsnGluIlePheIleProCys 236
 Db 750 ACTGAAGCTCAATGGCCCTAAGTCTCATAGAGTGTGAAGCAATATCACAAATCCATGC 809
 QY 237 ArgLeuAlaThrValAlaSerGlyAlaAlaSerGlyLysLeuLeuGluTyrGluValGly 256
 Db 810 AGGCTTGCTACTGTCGCTTCTGGAGCAGCTCTCGAAACCTTTTGCAGTATGAACCTTGC 869
 QY 257 GlyProArgValCysValGlnThrAlaTyrGlyValGluValGluValGluAsnPro 276
 Db 870 GGTCCTCGCTTTGCGTTCAAACAGCTTATGGTATAGAGGTTGAGTTGAAGCATATACC 929
 QY 277 TyrAspProAsnLeuMetValPheMetAspTyrArgAspTyrMetGlnGlnLysLeuGln 296
 Db 930 TATGATCCAAAGCCCTAATGGTTTTCATGATTATAGAGCTACACCAAAACATAATCTCAA 989

QY 297 CysSerGluGluGluTyrProThrPheLeuTyrValMetProMetSerProThrArgLeu 316
 Db 990 TCACATAGAGCACAAATATCCAAATTTTGTATGTATGTCATGCCATGTCTCCAACTAAGTA 1049
 QY 317 PhePheGluGluThrCysLeuAlaSerLysAspAlaMetProPheAspLeuLeuLysArg 336
 Db 1050 TTTCTTTGAGAAACTTGTGTGCTTCAAAAGAGGCATCCCTTTTGTAGTTATTGAAGACA 1109
 QY 337 LysLeuMetSerArgLeuLysThrLeuGlyIleGlnValThrLysValTyrGluGluGlu 356
 Db 1110 AAACATCATGTCAAGATTAAAGACTATGGGATCCGGAATTAACCAAACTTATGAAGAGGAA 1169
 QY 357 TrpSerTyrIleProValGlyGlySerLeuProAsnThrGluGlnLysAsnLeuAlaPhe 376
 Db 1170 TGGTCATATATCCAGTAGTGGATCCTTACCAAAATACCGAGCAAAAGAACCTTGATTT 1229
 QY 377 GlyAlaAlaSerMetValHisProAlaThrGlyTyrSerValValArgSerLeuSer 396
 Db 1230 GGTGCTGCTGTAGCATGGTGCATCCAGCCACAGGATATTCGGTTGTAGATCAGCTGCA 1289
 QY 397 GluAlaProLysTyrAlaSerValIleAlaLysIleLeuLysGlnAspAsnSerAlaTyr 416
 Db 1290 GAAGCTCCTAATATGACAGCAGTAATTCAAAGATTITAGGGAAGAGAAATTCAAACAG 1349
 QY 417 ValValSerGlyGlnSerSerAlaValAsnIleSerMetGlnAlaTrpSerSerLeuTrp 436
 Db 1350 ATGCTTGATCATGGAAGATACACACCAACATCTCAAGCAAGCTTGGGAAACACATTGG 1409
 QY 437 ProLysGluArgGlyArgGlnArgAlaPhePheLeuPheGlyLeuLeuIleValGln 456
 Db 1410 CCCCTTGAAGGAAAAAGACAGAGAGCATCTTCTCTTTGGATTAGACTGATTGTCCAG 1469
 QY 457 LeuAspIleGluAlaThrArgThrPhePheArgThrPhePheArgLeuProThrTrpMet 476
 Db 1470 ATGGATATGAGGGAGCCGACATCTTCCGACATTTCTTCCGCTTGCCCATCGATG 1529
 QY 477 TrpTrpGlyPheLeuGlySerSerLeuSerSerPheAspLeuValLeuPheSerMetTyr 496
 Db 1530 TGGTGGGGTTTCTTGGATCTTCTGTTATCATCAACTGACTTGATAATATTGGTGTTCAC 1589
 QY 497 MetPheValLeuAlaProAsnSerMetArgMetSerLeuValArgHisLeuLeuSerAsp 516
 Db 1590 ATGTTTATCATACAGCCGATAGCTAGCAATGGGTCTGTTAGACATTTGCTTCTGAC 1649
 QY 517 ProSerGlyAlaValMetValArgAlaTyrLeu 527
 Db 1650 CCGACAGGAGCAACAATGTTAAAGCGTATCTC 1682

RESULT 3

US-08-908-758-3
 ; Sequence 3, Application US/08908758
 ; Patent No. 6429356

GENERAL INFORMATION:

; APPLICANT: Calgene, Inc.
 ; TITLE OF INVENTION: METHODS FOR PRODUCING CAROTENOID COMPOUNDS AND SPECIALTY O
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Calgene, Inc.
 ; STREET: 1920 Fifth Street
 ; CITY: Davis
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 95616
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB
 ; COMPUTER: IBM PC
 ; OPERATING SYSTEM: Windows NT 4.0
 ; SOFTWARE: Microsoft Word For Windows 7.0a
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/908,758
 ; FILING DATE: 8-8-97
 ; CLASSIFICATION: 800

Db 667 AC GTTCTCTACGCCATCGCTATGACAAAGACAGAGTCTCTTTGAGGAGACATGTCTT 726
 Qy 324 AlaSerLysaspAlaMetProPheAspLeuLeuLysArgLysLeuMetSerArgLeuLys 343
 Db 727 GCTTCAAAGATGTCATGCCCTTTGATTGCTTAAAGAGAGCTCTTGTGAGATTAGAG 786
 Qy 344 ThrLeuGlyIleGlnValThrLysValTyrgLuGluGluTrpSerTyrlleProValGly 363
 Db 787 AC ACTCGGAATCGAATACTAAAGACTTACGAAGAGGAAATGGTCTTATATCCCCAGTAGT 846
 Qy 364 GlySerLeuProAsnThrGluGlnLysAsnLeuAlaPheGlyAlaAlaSerMetVal 383
 Db 847 GGTTCCTGCCAAACACGGAACAAAGAAATCTCGCCTTTGGTGTGCAGCTAGCATGGT 906
 Qy 384 HisProAlaThrGlyTyrsrValValArgSerLeuSerGluAlaProLysTyrlaAlaSer 403
 Db 907 CATCTGCAACAGGGTATTCAGTTGTGAGATCCTTGTCGAAGCTCCAAATACGATCA 966
 Qy 404 ValIleAlaLysIleLeuLysGlnAspAsnSerAlaTyrlValSerGlyGlnSerSer 423
 Db 967 GTCATCGGTAATATCTAAACATGACGACACTACTTCTCTCAACC-----AGACACATC 1020
 Qy 424 AlaValAsnIleSerMetGlnAlaTrpSerSerLeuTrpProLysGluArgLysArgGln 443
 Db 1021 AACACCAATATTCACGAAAGCTGGGATACTTTATGGCCACCAAGAAAGAACACAG 1080
 Qy 444 ArgAlaPhePheLeu 448
 Db 1081 AGAGCATCTCTTCTCA 1095

RESULT 4
 : Sequence 2, Application US/08908758
 : Patent No. 6423356
 : GENERAL INFORMATION:
 : APPLICANT: Calgene, Inc.
 : TITLE OF INVENTION: METHODS FOR PRODUCING CAROTENOID COMPOUNDS AND SPECIFIC
 : NUMBER OF SEQUENCES: 4
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Calgene, Inc.
 : STREET: 1920 Fifth Street
 : CITY: Davis
 : STATE: CA
 : COUNTRY: USA
 : ZIP: 95616
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB
 : COMPUTER: IBM PC
 : OPERATING SYSTEM: Windows NT 4.0
 : SOFTWARE: Microsoft Word For Windows 7.0a
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/908,758
 : FILING DATE: 8-8-97
 : CLASSIFICATION: 800
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 60/024,145
 : FILING DATE: 8-9-96
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Donna E. Scherer
 : REGISTRATION NUMBER: 34,719
 : NAME: Carl J. Schwedler
 : REGISTRATION NUMBER: 36,924
 : REFERENCE/DOCKET NUMBER: 122-1
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (916) 753-6313
 : TELEFAX: (916) 753-1510
 : INFORMATION FOR SEQ ID NO: 2 :
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 960 base pairs
 : TYPE: nucleic acid
 : STRANDEDNESS: double
 : TOPOLOGY: linear
 : MOLECULE TYPE: cDNA to mRNA

Db 430 TGGGTGGATGAATTTGAGGCTATGGACTTGTAGATTGTCTAGATGCTACCTGCTGGT 549
 Qy 169 ThrileValTyrLeuAspAsnAlaProValLeuIleGlyArgAlaTyrGlyArgVal 188
 Db 550 GCACAGTGTACATTGATGATAATGCGCTAAAGATCTTCATAGACCTTATGGAAGGTT 609
 Qy 189 SerArgHisLeuLeuHisGluGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 208
 Db 610 AACCGGAACAGCTGAATCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 669
 Qy 209 LeuAspSerLysValGluArgIleThrGluAlaGlyAspGlyHisSerLeuValValCys 228
 Db 670 CACCAAGCCCAAGTATAAGGTTGATTCATGAG---GAATCAAGTCAAGTCAAGTCAAG 726
 Qy 229 GluAsnGluIlePheIleProCysArgLeuAlaThrValAlaSerGlyAlaAlaSerGly 248
 Db 727 AATGATGGTATTACTATTGAGCAGCAGGCGGTGCTCGATGCAACATGGC---TTCTCTAGA 783
 Qy 249 LysLeuLeuGluTyrGluValGly---GlyProArgValCysValGlnThrAlaTyrGly 267
 Db 784 TCTCTTGTTCAGTATGATAAGCTTATAACCCCGG-----TATCAAGTTCCTATGCGC 837
 Qy 268 ValGluValGluValGluAsnProTyrAspProAsnLeuMetValPheMetAspTyr 287
 Db 838 ATTTGGCTGAAGTGAAGACACCCCTTTCATGTAACAGATGGTTTCATGATGG 897
 Qy 288 ArgAsp---TyrMetGlnLysLeuGlnCysSerGlu-----GluGluTyrProThr 304
 Db 898 CGAGATTCTCATTTGAAGAACAACTACTGATCTCAAGAGAGAGAAATAGTAGAATACCAACT 957
 Qy 305 PheLeuTyrValMetProMetSerProThrArgLeuPheGluGluThrCysLeuAla 324
 Db 958 TTTCTTATGCAATGCCATTTTCACCAAGGATATTTCTTGAAGAAACATCCTCGTA 1017
 Qy 325 SerLysAspAlaMetProPheAspLeuLeuLysArgLysLeuMetSerArgLeuLysThr 344
 Db 1018 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1077
 Qy 345 LeuGlyIleGlnValThrLysValTyrGluGluGluTyrPheValGlyGly 364
 Db 1078 TTGGGATAAAGTGAAGACATTTGAAGAGATGAATGTCTAATACCAATGGTGGT 1137
 Qy 365 SerLeuProAsnThrGluGlnLysAsnLeuAlaPheGlyAlaAlaAlaSerMetValHis 384
 Db 1138 CCACCTCCAGTATTACCTCAGAGAGCGTTGGAATCGGTGGTACAGCTGGCATGTTTAT 1197
 Qy 385 ProAlaThrGlyTyrSerValValArgSerLeuSerGluAlaProLysTyrAlaSerVal 404
 Db 1198 CCATCACCGTATATGTGCGCAAGGACACTAGCTGGCGGCTCCGTGTTGGCCAAATGCC 1257
 Qy 405 IleAlaLysIleLeuLysGlnAspAsnSerAlaTyrValValSerGlyGlnSerSerAla 424
 Db 1258 ATAATTCATACCTCGGTTCTGAAGAAGT-----CATTCGGGT 1296
 Qy 425 ValAsnIleSerMetGlnAlaTrpSerSerLeuTrpProLysGluArgGlnArg 444
 Db 1297 AATGAATATCACAGCTGTTTGAAGAGATTGTGGCTATAGAGAGAGAGAGAGAGAGAG 1356
 Qy 445 AlaPheLeuPheGlyLeuGluLeuIleValGlnLeuAspIleGluAlaThrArgThr 464
 Db 1357 GAGTTCTTCTGCTCGGTATGATATCTCTGAGGCTGATTACCTGCTGCTGCTGCTGCTGCT 1416
 Qy 465 PhePheArgThrPhePheArgLeuProThrTrpMetTrpTrpGlyPheLeuGlySerSer 484
 Db 1417 TTCCTTATGATGATCTTTTGAACCTTAGAACCTCGTTATGGCATGCTTCTATGCTCGA 1476
 Qy 485 LeuSerSerPheAspLeuValLeuPheSerMetTyrMetPheValLeuAlaProAsnSer 504
 Db 1477 TTGTTTCTACCTGACTCATAGTTTGGCTGCTCTCTATCTCTCTCATGCTCAATACT 1536
 Qy 505 MetArg-----MetSerLeuValArgHisLeu 513
 Db 1537 TCTAGATTTCAGATAATGACAAAGGAAGTGTTCATTAGTAATATGATCAACAATTTG 1596

Qy 514 LeuSerAsp 516
 Db 1597 TTACAGGAT 1605
 RESULT 6
 US-09-201-641-1
 ; Sequence 1, Application US/09201641A
 ; Patent No. 6232530
 ; GENERAL INFORMATION:
 ; APPLICANT: Cunningham Jr, Francis X
 ; APPLICANT: Dellapenna, Dean
 ; TITLE OF INVENTION: Method for Regulating Carotenoid Biosynthesis in
 ; TITLE OF INVENTION: Marigolds
 ; FILE REFERENCE: Quest 41-162
 ; CURRENT APPLICATION NUMBER: US/09/201,641A
 ; CURRENT FILING DATE: 1998-11-30
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: Patent In Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 1959
 ; TYPE: DNA
 ; ORGANISM: Tagetes erecta
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (304)..(1836)
 ; OTHER INFORMATION: beta-cyclase
 US-09-201-641-1
 Alignment Scores:
 Pred. No.: 6,31e-78 Length: 1959
 Score: 750.00 Matches: 180
 Percent Similarity: 55.85% Conservative: 97
 Best Local Similarity: 36.29% Mismatches: 189
 Query Match: 27.34% Indels: 30
 DB: 4 Gaps: 15
 US-09-701-395A-23 (1-529) x US-09-201-641-1 (1-1959)
 Qy 25 SerSerLysLeuAlaTyrAsnIleHisArgTyrGly-----SerSerCysArgValAsp 42
 Db 352 AGTAACAAATTTGCAGAAATTTGAACAATTTGAACAATTTGAACAATTTGAACAATTTGAACA 411
 Qy 43 PheGlnValArgAlaAspGlyGlySerGlySerArgSerValAlaTyrLysGluGly 62
 Db 412 TTTCAA-----GACTTTAGATTGGCCCAAAAATCC---CAATTCAAATTAGGG 459
 Qy 63 PheValAspGluGluAspPheIleLysAlaGlyGlySerGluLeuLeu---PheValGln 81
 Db 460 -----CAAAAATATGTTAAAGCTAGTAGTAGTGTGTTGTTAGAACTTGTCTCT 510
 Qy 82 MetGlnGlnThrLysSerMetGluLysGlnAlaLysLeuAlaAspLysLeuProIle 101
 Db 511 GAAATCAAGAAAGAAATCTTGATTTTGTATCTCTCTATGATGAT----- 555
 Qy 102 ProPheGlyGluSerValMetAspLeuValIleGlyCysGlyProAlaGlyLeuSer 121
 Db 556 CCATCAAGAAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 615
 Qy 122 LeuAlaAlaGluAlaLysLeuGlyLysValGlyLeuLeu-----GlyProAsp 139
 Db 616 GTGGCTCAACAAGTCTCTGAGGCTGCTCAGCTGCTCAATTTGACCCATCACCTAAA 675
 Qy 140 LeuProPheThrAsnAspTyrGlyValTrpGluAspGluPheLysAspLeuGlyLeuGlu 159
 Db 676 CTCATTGGCCCAATATATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 735
 Qy 160 ArgCysIleGluHisAlaTrpLysAspThrIleValTyrLeuAspAsnAlaProVal 179
 Db 736 CATTTGTTGGATACAACTTGTGCAAGTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 795
 Qy 180 LeuIleGlyArgAlaTyrGlyArgValSerArgHisLeuLeuHisGluGluLeuLys 199

NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Calgene, Inc.
STREET: 1920 Fifth Street
CITY: Davis
STATE: CA
COUNTRY: USA
ZIP: 95616

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB
COMPUTER: IBM PC
OPERATING SYSTEM: Windows NT 4.0
SOFTWARE: Microsoft Word For Windows 7.0a
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/908,758
FILING DATE: 8-8-97
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
PRIOR APPLICATION NUMBER: 60/024,145
FILING DATE: 8-9-96
ATTORNEY/AGENT INFORMATION:
NAME: Donna E. Scherer
REGISTRATION NUMBER: 34,719
NAME: Carl J. Schwedler
REGISTRATION NUMBER: 36,924
REFERENCE/DOCKET NUMBER: 122-1
TELEPHONE: (916) 753-6313
TELEFAX: (916) 753-1510
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1590 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
US-08-908-758-4

Alignment Scores:
Pred. No.: 8,72e-77
Score: 739.00
Percent Similarity: 56.70%
Best Local Similarity: 36.61%
Query Match: 26.94%
DB: 4

US-09-701-395A-23 (1-529) x US-08-908-758-4 (1-1590)

QY 107 ValMetAspLeuValValleGlySerLeuAlaGlyLeuSerLeuAlaGluAla 126
DB 1304 GTTGTGATCTAGCTATCTGTCTGCTCATGCTCCCTCCCAAACTCATTTGGCCTAAC 1185
QY 127 AlAlYSLeuGlyLeuLysValGlyLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 144
DB 1244 TCGGAAGCTGGAGTCTCTGTCTGCTCATGCTCCCTCCCAAACTCATTTGGCCTAAC 1185
QY 145 AsnTyrGlyValTyrGluAspGluPheLysAspLeuGlyLeuGluArgCysLeuHls 164
DB 1184 AACTACGGAGTGTGGTGTGAGGAGTTCGAGGAGTTCGAGGAGTTCGAGGAGTTC 1125
QY 165 AlarTyrAspThrLeuValTyrLeuAspAsnAlaPheValLeuLeuLeuLeuLeu 184
DB 1124 ACCTGGTCCGGCGGCTGCTCTATCATCAACGAGCTCCCAAAAGGACCTGTCGCGGCT 1065
QY 185 TyrGlyArgValSerArgHisLeuLeuHisGluGluLeuLeuLysArgCysValGluSer 204
DB 1064 TACGGGAGAGTGAACCGTAACAGCTCAAAATCCAAAGATGCTTCAGAGTGCATCACCAC 1005
QY 205 GlyValSerTyrLeuAspSerLysValGluArgLleThrGluAlaGlyAspGlyHisSer 224
DB 1004 GGTGTGTAGTTCATCAGGCTAAAGTCACTGACGTGGTGGTTCAGGAGTTCAGGAGTTC 948
QY 225 LeuValValCysGluAsnGluLeuPheLeuProCysArgLeuAlaThrValAlaSerGly 244

796 AGTCTTAATAGACCATATGCAAGAGTCATAGAAACAACCTTAAACAAGAGATGTTACAA 855
QY 200 ArgCysValGluSerGlyValSerTyrLeuAspSerLysValGluArgLleThrGluAla 219
DB 856 AAGTGTATAGCAATGGGTAAAGTTTATCATCAAGCAAAAGTCATCAAAAGTGCATGAA 915
QY 220 GlyAspGlyHisSerLeuValValCysGluAsnGluLeuPheLleProCysArgLeuAla 239
DB 916 --GAGTAAATCTTGTGTGATTTCTAATGATGTTGTCACATGATCAAGCCACTTTGTT 972
QY 240 ThrValAlaSerGlyAlaLaSerGlyLysLeuLeuGluTyrGluValGly---GlyPro 258
DB 973 CTTCATCAACTGCT---TTTTCAGAGCTTTAGTTCAATGATGATAAGCTTATAACCT 1029
QY 259 ArgValCysValGlnThrAlaTyrGlyValGluValGluValGluValGluValGlu 278
DB 1030 GGG-----TACCAAGTGGCTTATGGATTTTATGCGGAGTTGAAGAACACCTTTTGAC 1083
QY 279 ProAsnLeuMetValPheMetAspTyrArgAsp---TyrMetGlnGlnLysLeuGln--- 296
DB 1084 GTTGATAAATCTGTTTATGATGGATGGAGAGTTCACACCTTGATCAAAATCTTGAAT 1143
QY 297 ---CysSerGluGluTyrProThrPheLeuTyrValMetProMetSerProThrArg 315
DB 1144 AAAGTAGAATTCAGAAATCCAACTTTTATACGGGATGCGCATTTCTGCTACAGA 1203
QY 316 LeuPheGluGluThrCysLeuAlaSerLysAspAlaMetProPheAspLeuLeuLys 335
DB 1204 ATCTTTCTTGAAGAAACATCATCTGCTGCTCCGGGTTGAAGATGGAAGATATCAA 1263
QY 336 ArgLysLeuMetSerArgLeuLysThrLeuGlyLleGlnValThrLysValTyrGluGlu 355
DB 1264 GAAAGAATGCTTACAGGCTAAAGCATTTGGGATGATAAAGTAAAGCATTCGAAGAAG 1323
QY 356 GluTrpSerTyrIleProValGlyLysLeuProAsnThrGluGlnLysAsnLeuAla 375
DB 1324 GAACGTGTGTTATCCGATGGGGGCGCCCTACCAGTGTCTCCCTCAACGGGTCTTGA 1383
QY 376 PheGlyAlaAlaSerMetValHisProAlaThrGlyTyrSerValValArgSerLeu 395
DB 1384 ATAGTGTGACAGAGAAATGTCATCCGTCACCGATACATGTTGGTGGCAAGAGCTA 1443
QY 396 SerGluAlaProLysTyrAlaSerValIleAlaLysLleLeuLysGlnAspAsnSerAla 415
DB 1444 GCAGCGCGCCGATGTTGCAAGTCAATAATCCGGTATCTATAACGAAAAAGTATG 1503
QY 416 TyrVal---ValSerGlyLysSerSerAlaValaSerMetGlnAlaTrpSerSer 434
DB 1504 GTGGCGGAGTCCACGGAGATGATTACAGCCGGAATA-----TGGAGAGAA 1551
QY 435 LeuTyrProLysGluArgLysArgGlnArgAlaPheLeuPheGlyLeuLeuLeuLeu 454
DB 1552 TTGTGCTTATGAAAGAGGAGACAAAGAGGAGTGTGTTGTTGGGATGGATATATG 1611
QY 455 ValGlnLeuAspLleGluAlaThrArgThrPhePheArgThrPhePheArgLeuProThr 474
DB 1612 TTGAAGCTGATTTGGAAGTACTAAGAGTCTTTGATGCGCTTTTTCGACTTGAACCT 1671
QY 475 TrpMetTyrPheGlyPheLeuGlySerSerLeuSerSerPheAspValLeuPheSer 494
DB 1672 CGTTATTGGCATGGTTTTTGTGCTGAGGTTGTTTCTACCGGAGTATGACGTTTGG 1731
QY 495 MetTyrMetPheValLeuAlaProAsnSerMetArgMetSerLeuVal 510
DB 1732 CTATCGCTTTTTCGCTATCGTTCGAATACITGTAGATGTTGAAATATG 1779

RESULT 7
US-08-908-758-4/c
Sequence 4, Application US/08908758
Patent No. 6429356
GENERAL INFORMATION:
APPLICANT: Calgene, Inc.
TITLE OF INVENTION: METHODS FOR PRODUCING CAROTENOID COMPOUNDS AND SPECIALTY OIL

```

Db 947 ACTGTTGCTCAGTGGTGAAGATTCAAGCTTCTGTCTTCTCAGCCTACTGCT 888
Qy 245 AlaAlaSerGlyLysLeuLeuGluThrGluValGly--GlyProArgValCysValGln 263
Db 887 ---TTTCAAGATGCTTGGTTCAGTATGATAGCCTTATAACCCCTGGG-----TATCAA 837
Qy 264 ThrAlaTyrGlyValGluValGluValGluValGluValGluValGluValGluVal 283
Db 836 GTAGCTTAGTATGATCTCTGAGTGGTATGCTACCCGTTTGTGATGATGATGATG 777
Qy 284 PheMetAspTyrArgAsp-----TyrMetGlnGlnLysLeuGlnCysSer 298
Db 776 TTAGTGGCTGAGAGATGAAGCATCTTCACCGCTACCTGAGGTTAAGAGCGGAACAG 717
Qy 299 GluGluGluThrProThrPheLeuTyrValMetProMetSerProThrArgLeuPhe 318
Db 716 -----AAGATCCCTACCTTCTGTACCGGATGCCGTTTCTTCCAACAGATCTTCT 653
Qy 319 GluGluThrCysLeuAlaSerLysAspAlaMetProPheAspLeuLeuLysArgLys 338
Db 662 GAAGAGAGCTCTGTGGCTAGCGCGGTCTGAAGATGCAAGATATCCAAGAGAGATG 603
Qy 339 MetSerArgLeuLysThrLeuGluValGlnValThrLysValTyrGluGluThrPhe 358
Db 602 GTTCGAGGCTGAACATTTGGGATCAACGTCGAGGATTCAGGATTCAGGATTCAGG 543
Qy 359 TyrIleProValGlyGlySerLeuProAsnThrGluGlnLysAsnLeuAlaPheCly 378
Db 542 GTGATCCCTATGAGGCTCTTACCGGCTTTCGCTCAAGAGTGTGGGATTTGGCGGT 483
Qy 379 AlaAlaSerMetValHisProAlaThrGlyTyrSerValValArgSerLeuSerGlu 398
Db 482 ACGCGGGGATGTTTCATCTTCGACTGTTACATGGTGTCTAGGATCTTTCAGCTG 423
Qy 399 ProLysTyrAlaSerValIleAlaLysLysLeuLysGlnAspAsnSerAlaTyrVal 418
Db 422 CCGATAGTTGCAACGCTATAGTCGG-----AsnIleSerMetGlnAlaTyrSer 435
Qy 419 SerGlyGlnSerSerAlaVal-----AsnIleSerMetGlnAlaTyrSerSer 435
Db 386 GGTGTAACAACAACGCTTGAGAGGATGAGCTCTCGCTGAGGTGGAGAGACTTG 327
Qy 436 TrpProLysGluArgLysAlaGlnArgAlaPhePheLeuPheGlyLeuGluLeuVal 455
Db 326 TGGCTATTGAGCGCGGACAGAGAGGAGTCTTCTGTTTGGGATGATATTGCTG 267
Qy 456 GlnLeuAspIleGluAlaThrArgThrPhePheArgThrPhePheArgLeuProThr 475
Db 266 AAGCTTGATTTGATGCTACTAGGAGTCTTTGATGCTTCTTGACCTGGAACCG 207
Qy 476 MetTrpTrpGlyPheLeuGlySerSerLeuSerSerPheAspLeuValLeuPheSer 495
Db 206 TACTGCGATGTTTCTTCTGCTGAGGCTGTTTCTCCCGGACCTGCTTCTTCGG 147
Qy 496 TyrMetPheValLeuAlaProAsnSerMetMetSerLeuVal-----510
Db 146 TCGCTCTTCCATGCTCTTAATACCTCGAGGTTGGAGATCATGCAAGGAAGTGT 87
Qy 510 -----510
Db 86 CTTCTGTGAATGATCAACAATTTGGTACAGATAGAGCGCAATTCAGCACACTG 27
Qy 511 ArgHisLeuLeuSerAspProSer 518
Db 26 CGGCGGTACTAGTGGATCCGAGC 3

```

RESULT 8
 US-08-399-561-4
 ; Sequence 4, Application US/08399561
 ; Patent No. 5792903
 ; GENERAL INFORMATION:
 ; APPLICANT: Hirschberg, Joseph

```

APPLICANT: Cunningham Jr., Francis X.
APPLICANT: Gantt, Elisabeth
TITLE OF INVENTION: Lycopene Cyclase Gene
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kohn & Associates
STREET: 30500 No. 5792903thwestern Highway, Suite 410
CITY: Farmington Hills
STATE: MI
COUNTRY: US
ZIP: 48334
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/399,561
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kohn, Kenneth I.
REGISTRATION NUMBER: 30,955
REFERENCE/DOCKET NUMBER: P-305 (Hebrew Univ.)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 810-539-5050
TELEFAX: 810-539-5055
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1614 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: Nicotiana tabacum
STRAIN: cv. xanthii
TISSUE TYPE: Leaf
US-08-399-561-4

Alignment Scores:
Pred. No.: 1,54e-76 Length: 1614
Score: 737.00 Matches: 174
Percent Similarity: 54.70% Conservative: 174
Best Local Similarity: 33.40% Mismatches: 180
Query Match: 26.87% Indels: 56
DB: 1 Gaps: 15

US-09-701-395A-23 (1-529) x US-08-399-561-4 (1-1614)
Qy 14 ProValTyrThrPheGlyThrArgAsnLeuSerSerLysLeuAlaTyrAsnIleHis 33
Db 106 CCAGTTCATGATTTCTGTTAAA-----GCTAGCTCCTTTAACTCTGTAAGCCCCAT 159
Qy 34 ArgTyrGlySerSerCysArgValAspPheGlnValArgAlaAspGlyGlySerGlySer 53
Db 160 AAGTTTGGTCTTAGG-----AAATTTGTGAAATTTGGGTAAGGGTT 204
Qy 54 ArgSerSerValAlaTyrLysGluGlyPheValAspGluAspPheIleLysAlaGly 73
Db 205 -----TGTTTAAAGCTAAG 219
Qy 74 GlySerGluLeuLeu-----PheValGlnMetGlnGlnThrLysSerMetGluLysGlnAla 92
Db 220 AGTAGTGCCCTTTTGGAGCTTGCTACCTGACACCAAGGAAATCTTCATTTGAGCTT 279
Qy 93 LysLeuAlaAspLysLeuProPheGlyGlySerValMetAspLeuValVal 112
Db 280 CCTATGATGAC-----CCTTCAAAAGGCTTGTGTAGATCTAGCTG 324
Qy 113 IleGlyCysGlyProAlaGlyLeuSerLeuAlaAlaGluAlaLysLeuGlyLeuLys 132
Db 325 GTTGTGTGGTGGACCGCTGGACTTGCAGTTCACAGCAGGTTTCGGAGGCTGGACTATCG 384

```

Qy	487	SerPheAspLeuValLeuPheSerMetTyrMetPheValLeuAlaProAsnSerMetArg	506
Db	1432	CTTCTGAGCTTATATTTTCGGGCTGTCCCTTTTCTCTCGCGCTTCAAATACTTCTTAGA	1491
Qy	507	-----MetSerLeuValArgHisLeuLeuSer	515
Db	1492	ATAGAGATTATGACAAAGGAAGCTCTTCTTTGGTAAATATGATCAACAAATTTGTTACAG	1551
Qy	516	Asp 516	
Db	1552	GAT 1554	
RESULT 9			
US-09-134-607A-11			
; Sequence 11, Application US/09134607A			
; Patent No. 6252141			
; GENERAL INFORMATION:			
; APPLICANT: Joseph Hirschberg et al.			
; TITLE OF INVENTION: POLYNUCLEOTIDES CONTROLLING THE EXPRESSION			
; TITLE OF INVENTION: OF AND CODING FOR GENE B IN TOMATO AND USE			
; TITLE OF INVENTION: OF SAME FOR ALTERING CAROTENOID			
; TITLE OF INVENTION: BIOSYNTHESIS			
; NUMBER OF SEQUENCES: 25			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: Mark M. Friedman c/o Anthony Castorina			
; STREET: 2001 Jefferson Davis Highway, Suite 207			
; CITY: Arlington			
; STATE: Virginia			
; COUNTRY: United States of America			
; ZIP: 22202			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk			
; COMPUTER: Twinhead, Slimnote 890TX			
; OPERATING SYSTEM: MS DOS version 6.2,			
; OPERATING SYSTEM: Windows version 3.11			
; SOFTWARE: Word for Windows version 2.0,			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/09/134,607A			
; FILING DATE:			
; CLASSIFICATION: 800			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER:			
; FILING DATE:			
; ATTORNEY/AGENT INFORMATION:			
; NAME: Friedman, Mark M.			
; REGISTRATION NUMBER: 33,883			
; REFERENCE/DOCKET NUMBER: 325/12			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: 972-3-562553			
; TELEFAX: 972-3-562554			
; TELEX:			
; INFORMATION FOR SEQ ID NO: 11:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 2897			
; TYPE: nucleic acid			
; STRANDEDNESS: double			
; TOPOLOGY: linear			
US-09-134-607A-11			
Alignment Scores:			
Pred. No.: 2,08e-74 Length: 2897			
Score: 722.50 Matches: 165			
Percent Similarity: 58.94% Conservative: 79			
Best Local Similarity: 39.86% Mismatches: 151			
Query Match: 26.34% Indels: 19			
DB: 4 Gaps: 9			
US-09-701-395A-23 (1-529) x US-09-134-607A-11 (1-2897)			
Qy	104	GlyClnSerValMetAspLeuValValIleGlyCysGlyProAlaGlyLeuSerLeuAla	123
Db	1463	GGTCGGCGCTCAATTCGACGCTGATCATATTATCGGAGCTGGCCCTCGGCTTCAGGTAGCT	1522

124	AlaGluAlaIalaLysLeuGlyLeuLysValGlyLeuIleGlyProAspLeuProPheThr	143
QY	:::::	
1523	GAACAAGTTTCTAAATTAAGGATAATTGGTGTCACCCCTTCA--CCACTCTCC	1579
Db	:::::	
144	-----AsnAsnTyrglyValTrpGluAspGluPhelysAspLeuGlyLeuGluArg	160
QY		
1580	ATGTGCCAATATTAATATGGTGGTGTGATGAGTTTGAGAAATTTAGGACGTGGAAGAT	1639
Db		
161	CysIleGluHisAlaTrpLysAspThrIleValTyrlauAspAsnAlaProValLeu	180
QY		
1640	TGTTTTAGATCATAAATGCCTTACITGTGTGCATATAATGATACAAGACTAAGTAT	1699
Db		
181	IleGlyArgAlaTyrglyArgValSerArgHisLeuLeuHisGluGluLeuLysArg	200
QY		
1700	TTGGGAAGACCATAATGGTAGAGTTAGTAGAAGAAGCAGTGAATTTGAAATTTTGAACAGT	1759
Db		
201	CysValGluSerGlyValserTyrlauAspSerLysValGluArgIleThrGluAlaGly	220
QY		
1760	TGTGTTGAGACAGAGTGAAGTTTATAAGCTAAGGTTTGGAAAGTG---GAACATGAA	1816
Db		
221	AspGlyHisSerLeuValValCysGlnAsnGluIlePheIleProCysArgLeuAlaThr	240
QY	:::::	
1817	GAATTTGAGTCTCAATGTTTGTGATGATGCGTAAGAAGATAAGAGCTAGTTGGTTGG	1876
Db		
241	ValAlaSerGlyAlaIalaSerGlyLysLeuLeuGluTyrgluValGlyProArg---	259
QY		
1877	GATGCAAGTGGTTTGTCTAGT---GATTTATAGAGTATGAC-----AAGCCAAGAAAC	1927
Db		
260	ValCysValGlnThrAlaTyrglyValGluValGluValGluAsnAsnProTyrraspPro	279
QY		
1928	CATGTTTATCAAATTCATCGGCTTTAGTAGAAGTTGATATCATTCATTGTTGATTG	1987
Db		
280	AsnLeuMetValPheMetAspTyrrArgAspTyrr-----MetGlnLynLysLeuGln	296
QY	:::::	
1988	GATAAAATGGTGCTTATGCGATTGGAGGATTCCTCATTTAGTAAATGAGCCATATTAAAG	2047
Db		
297	CysSerGluGluGluTyrrProThrPheLeuTyrrValMetProMetSerProThrArgLeu	316
QY	:::::	
2048	GTGAATAATGCTCAAGAACCACCAATCTGTATGCAATGCCATTTGATAGAAATTTGGTT	2107
Db		
317	PhePheGluGluThrCysLeuAlaSerLysAspAlaMetProPheAspLeuLeuLysArg	336
QY		
2108	TTCTTTGGHAGAGACTTCTTGGTGAGTCGCTGTTATCGTATATGGAAGTAAAGA	2167
Db		
337	LysLeuMetSerArgLeuLysThrLeuGlyIleGlnValThrLysValTyrgluGluGlu	356
QY	:::::	
2168	AGGATGTTGGCAAGATTAAGGCATTTGGGATCAAAAGTGAGAAAGTGTTATTGAGGAAGAG	2227
Db		
357	TrpSerTyrlleProValGlyGlySerLeuproAsnThrGluGlnLysAsnLeuAlaphe	376
QY		
2228	AAATGTGTCATCCCTATGGGAGGACCACTTCCGCGGATTCCTCAAAATGTTATGGCTATT	2287
Db		
377	GlyAlaAlaAlaSerMetValHisProAlaThrGlyTyrrSerValValArgSerLeuSer	396
QY	:::::	
2288	GGTGGGAATTCAGGATAGTTTCATCCATCAACGGGTACATGTGGCTAGGAGCATGGCT	2347
Ddb		
397	GluAlaProLysTyrrAlaSerValIleAlaLysIleLeuLysGlnAspAsnSerAlaTyrr	416
QY		
2348	TTAGCACCACTACTAGCTGAAGCATCTGTCGAGGGCTT-----GGCTCAACAAGA	2398
Ddb		
417	ValValSerGlyGlnSerSerAlaValAsnIleserMetGlnAlatrpSerLeutrp	436
QY	:::::	
2399	ATGATAAGGGTCT-----CAACTTTACCATAGACTTGTGGAATGGTTGGTGG	2446
Db		
437	ProLysGluArgLysGlnArgAlaPhePheLeupheGlyLeuGluLeuIleValGln	456
QY		
2447	CCTTTGGATAGAAGATGTTTAGAGAAATGTTATTCATTTGGGATGAGACATTTGTTGAAG	2506
Ddb		
457	LeuAspileGluAlaThrArgThrPhePheArgThrPhePheargLeuproThrIrpmet	476
QY		
2507	CTTGATTTGAAGGCATAGAGATGTTTGACGCTTTCTTGTATCTTGATCCTAAAATAC	2566
Ddb		
477	TrpTrpGlyPheLeuGlySerSerLeuSerSerPheAspLeuValLeuPheSerMetTyrr	496
QY		

```

Db      2567 TGGCAGGGTTCCTTCTTCAAGATGCTCTCAAGACATCTGGTTACTCAGCTTGCT 2626
QY      497 MetpheValLeuAlaProAsnSerMetArgMetSerLeuVal 510
Db      2627 CTTTCGGCAGTGCGCTCAAAATTTGACTAGGTTGGATATTGTT 2668

RESULT 10
US-09-134-607A-15
; Sequence 15, Application US/09134607A
; Patent No. 6252141
; GENERAL INFORMATION:
; APPLICANT: Joseph Hirschberg et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES CONTROLLING THE EXPRESSION
; OF AND CODING FOR GENE B IN TOMATO AND USE
; TITLE OF INVENTION: OF SAME FOR ALTERING CAROTENOID
; TITLE OF INVENTION: BIOSYNTHESIS
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Mark M. Friedman c/o Anthony Castorina
; STREET: 20001 Jefferson Davis Highway, Suite 207
; CITY: Arlington
; STATE: Virginia
; COUNTRY: United States of America
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk
; COMPUTER: Twinhead, Slimnote 890TX
; OPERATING SYSTEM: MS DOS version 6.2,
; OPERATING SYSTEM: Windows version 3.11
; SOFTWARE: Word for Windows version 2.0,
; CURRENT APPLICATION DATA: US/09/134,607A
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Friedman, Mark M.
; REGISTRATION NUMBER: 33,883
; REFERENCE/DOCKET NUMBER: 325/12
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 972-3-562553
; TELEFAX: 972-3-562554
; TELEX:

```

```

Db 2995 CTTTTCGACATGCTCAAAATTGCATAGGTGGATATGTT 3036

RESULT 11
US-09-134-607A-8
; Sequence 8, Application US/09134607A
; Patent No. 6252141
; GENERAL INFORMATION:
; APPLICANT: Joseph Hirschberg et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES CONTROLLING THE EXPRESSION
; TITLE OF INVENTION: OF AND CODING FOR GENE B IN TOMATO AND USE
; TITLE OF INVENTION: OF SAME FOR ALTERING CAROTENOID
; TITLE OF INVENTION: BIOSYNTHESIS
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Mark M. Friedman c/o Anthony Castorina
; STREET: 20001 Jefferson Davis Highway, Suite 207
; CITY: Arlington
; STATE: Virginia
; COUNTRY: United States of America
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk
; COMPUTER: Twinhead, Slimnote 890TX
; OPERATING SYSTEM: MS DOS version 6.2,
; OPERATING SYSTEM: Windows version 3.11
; SOFTWARE: Word for Windows version 2.0,
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/134,607A
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Friedman, Mark M.
; REGISTRATION NUMBER: 33,883
; REFERENCE/DOCKET NUMBER: 325/12
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 972-3-5625553
; TELEFAX: 972-3-5625554
; TELEX:
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1666
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-09-134-607A-8

Alignment Scores:
Pred. No.: 1-85e-74 Length: 1666
Score: 719,50 Matches: 164
Percent Similarity: 59.17% Conservative: 78
Best Local Similarity: 40.10% Mismatches: 148
Query Match: 26.23% Indels: 19
DB: 4 Gaps: 9

US-09-701-395A-23 (1-529) x US-09-134-607A-8 (1-1666)
Qy 109 AspLeuValIleGlyCysGlyProAlaGlyLeuSerLeuAlaGluAlaLaLaLa
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 247 GACGTGATCATATTGCGAGCTGGCCCTGCTGGGCTCAGGCTAGCTGAACAAGTTCTCTA
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 129 LeuGlyLeuLysValGlyLeuIleGlyProAspLeuProPheThr-----AspAla
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 307 TATGGTATTAAAGTATGTTGTGTGACCCCTTCA---CCACTCTCCATGTGGCCAAATA
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 146 TyrGlyValTrpGluAspGluPheLysAspLeuGlyLeuGluArgCysrileGluHisAla
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 364 TATGCTGTTGGGTGTGATGAGTTTGAGAAATTAGGACTGGAAATGTTTAGATCATCAT
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 166 TrpLysAspThrIleValTyrLeuAspAsnAspAlaProValLeuIleGlyArgAlaTyr
||| ||| |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

Db 424 TGGCCTATGACCTTGTGTCATATAAATGATACAAACAACTAAGTATTTCGGGAAGACCATAT 483
Qy 186 GlyArgValSerArgHisLeuLeuHisGluGluLeuLysArgCysValGluSerGly 205
Db 484 GGTAGAGTTAGTAAAGAAAGCTGAAGTTGAATTTGTAATGTTGGTTGAGAACAGA 543
Qy 206 ValSerTyrLeuAspSerLysValGluArgIleThrGluAlaGlyAspGlyHisSerLeu 225
Db 544 GTGAAGTTTATAAGCTTAAGCTTTGGAAGCTG---GAACATGAAGAATTTGAGTCTCA 600
Qy 226 ValValCysGluAsnGluIlePheIleProCysArgLeuAlaThrValAlaSerGlyAla 245
Db 601 ATTGTTTGTGATGAGTGAAGAGATAGAGTAGTTGGTTGGATGCAAGTGGTTT 660
Qy 246 AlaSerGlyLysLeuLeuGluTyrGluValGlyGlyProArg---ValCysValGlnThr 264
Db 661 GCTAGT---GATTTTATAGAGTATGAC-----AGGCCAAGAACCATGGTTATCAAAAT 711
Qy 265 AlaTyrGlyValGluValGluValGluAsnProTyrAspProAsnLeuMetValPhe 284
Db 712 GCTATGGGGTTTGTAGTAAAGTTGATAATCATCCATTGATTTGGATATAAATGGTGCTT 771
Qy 285 MetAspTyrArgAspTyr-----MetGlnGlnLysLeuGlnCysSerGluGluGlu 301
Db 772 ATGGATTGGAGGATTCATTTGGTAAATGAGCCATATTTAAGGTGAATTAAGCTAAA 831
Qy 302 TyrProThrPheLeuTyrValMetProMetSerProThrArgLeuPhePheGluGluThr 321
Db 832 GAACCAACATCTTGTATGCAATGCCATTTGATAGAGATTTGGTTTCTTGGAGAGACT 891
Qy 322 CysLeuAlaSerLysAspAlaMetProPheAspLeuLeuLysArgLysLeuMetSerArg 341
Db 892 TCTTTGGTGGTCCGCTGTTTATCGTATATGAAGTAAAGAAAGAGGATGGTGCAAGA 951
Qy 342 LeuLysThrLeuGlyIleGlnValThrLysValTyrGluGluGluTyrSerTyrIlePro 361
Db 952 TTAAGGCATTTGGGATCAAGTGAAGTGTATTGAGGAGAGAAATGTGTATCCCT 1011
Qy 362 ValGlyGlySerLeuProAsnThrGluGlnLysAsnLeuAlaPheGlyAlaAlaSer 381
Db 1012 ATGGGAGGACCACTCCGCGGATCTCCAAAATGTTATGGCTATTGGTGGGAATTCAGG 1071
Qy 382 MetValHisProAlaThrGlyTyrSerValValArgSerLeuSerGluAlaProLysTyr 401
Db 1072 ATAGTTTATCCATCAACAGGTACATGGTGGCTAGGAGCATGGCTTTAGCACCACTACTA 1131
Qy 402 AlaSerValIleAlaLysIleLeuLysGlnAspAsnSerAlaTyrValValSerGlyGln 421
Db 1132 GCTGAAGCCATCGTCGAGGGGCTT-----GGCTCAACAAGATGATAGAGGGTCT 1182
Qy 422 SerSerAlaValAsnIleSerMetGlnAlaTrpSerSerLeuTrpProLysGluArgLys 441
Db 1183 -----CAACTTTTACCATAGATTTGGAATGTTGTGGCTTTGGATAGAAGA 1230
Qy 442 ArgGlnArgAlaPhePheLeuPheGlyLeuLeuIleValGlnLeuAspIleGluAla 461
Db 1231 TGTGTTAGAGATTTTATTCATTTGGGATGGAGACATGTTTGAAGCTGTATTTGANAAGG 1290
Qy 462 ThrArgThrPhePheArgThrPhePheArgLeuProThrTrpMetTrpTrpPheLeu 481
Db 1291 ACTAGGAGATTTGTTGAGCTTTCTTTGATCTTATCCCTAAATACTGGCAAGGGTCTCT 1350
Qy 482 GlySerSerLeuSerSerPheAspLeuValLeuPheSerMetTyrMetPheValLeuAla 501
Db 1351 TCTTCAAGATGTCTGTCAAAGAACTTTGGTTTACTACGCTGTGTCTCTTTCGGACATGGC 1410
Qy 502 ProAsnSerMetArgMetSerLeuVal 510
Db 1411 TCACACATGACTAGTGGATATGT 1437

RESULT 12

US-09-134-607A-13

; Sequence 13, Application US/09134607A

; Patent No. 6252141
; GENERAL INFORMATION:
; APPLICANT: Joseph Hirschberg et al.
; TITLE OF INVENTION: POLYNUCLEOTIDES CONTROLLING THE EXPRESSION
; OF AND CODING FOR GENE B IN TOMATO AND USE
; TITLE OF INVENTION: OF SAME FOR ALTERING CAROTENOID
; TITLE OF INVENTION: BIOSYNTHESIS
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Mark M. Friedman c/o Anthony Castorina
; STREET: 20001 Jefferson Davis Highway, Suite 207
; CITY: Arlington
; STATE: Virginia
; COUNTRY: United States of America
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk
; COMPUTER: Twinhead, Slimnote 890TX
; OPERATING SYSTEM: MS DOS version 6.2,
; SOFTWARE: Word for Windows version 2.0,
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/134,607A
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Friedman, Mark M.
; REGISTRATION NUMBER: 33,883
; REFERENCE/DOCKET NUMBER: 325/12
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 972-3-5625553
; TELEFAX: 972-3-5625554
; TELEX:
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1666
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-09-134-607A-13
Alignment Scores:
Pred. No.: 1.85e-74 Length: 1666
Score: 719.50 Matches: 164
Percent Similarity: 59.17% Conservative: 78
Best Local Similarity: 40.10% Mismatches: 148
Query Match: 26.23% Indels: 19
DB: 4 Gaps: 9
US-09-701-395A-23 (1-529) x US-09-134-607A-13 (1-1666)
Qy 109 AspLeuValValIleGlyCysGlyProAlaGlyLeuSerLeuAlaGluAlaLys 128
Db 247 GACGTGATCATATATCGGAGCTGGCCCTGCTGAGCTAGCTGACAAAGTTTCTAAA 306
Qy 129 LeuGlyLeuLysValGlyLeuIleGlyProAspLeuProPheThr-----AsnAsn 145
Db 307 TATGGTATTAGGTATGTTGTTGACCCCTTCA---CCACTCTCCATGTGGCCAAATAAT 363
Qy 146 TyrGlyValTrpGluAspGluPheLysAspLeuGlyLeuGluArgCysIleGluHisAla 165
Db 364 TATGGTGTGGGTGTGATGAGTTTGAGATTTAGACTGGAAATTTGTTAGATCAAAA 423
Qy 166 TrpLysAspThrIleValTyrLeuAspAsnAlaProValLeuIleGlyArgAlaTyr 185
Db 424 TGGCCTATGACTTGTGTGCATATAAATGATAACAAACTAAGTATTGGGAAGACCATAT 483
Qy 186 GlyArgValSerArgHisLeuLeuHisGluGluLeuLysArgCysValGluSerGly 205
Db 484 GGTAGAGTTAGTAAAGAAAGCTGAAGTTGAATTTGTTGAATAGTGTGTGAGAACAGA 543

QY 226 ValValCysGluAsnGluPheIleProCysArgLeuAlaThrValAlaSerGlyAla 245
 Db 1811 ATTTGTTGTGATGCTAGACAGATAGAGGTAGTTTGGTTGGATCGAAGTGGTTT 1870
 QY 246 AlaSerGlyLysLeuGluTyrGluValGlyGlyProArg---ValCysValGlnThr 264
 Db 1871 GCTAGT--GATTTATAGAGTATGAC-----AGGCCAAGAACCACCATGGTTATCAATT 1921
 QY 265 AlaTyrGlyValGluValGluValGluAsnAsnProTyrAspProAsnLeuMetValPhe 284
 Db 1922 GCTCATGGGGTTTGTAGCAAGCTGATATCATCCATTTGATTTGGATAAATGGTGCTT 1981
 QY 285 MetAspTyrArgAspTyr-----MetGlnGlnLysLeuGlnCysSerGluGluGln 301
 Db 1982 ATGGATTGAGGGATTCATTTGGGTAATGACCATATTTAGGGTGAATATGCGTAA 2041
 QY 302 TyrProThrPheLeuTyrValMetProMetSerProThrArgLeuPheGluThr 321
 Db 2042 GAACCAACATTCCTTTATGCAATGCCATTCATAGAGATTTGGTTTCTTGAAGAGACT 2101
 QY 322 CysLeuAlaSerLysAspAlaMetProPheAspLeuLysArgLysLeuMetSerArg 341
 Db 2102 TCTTTGGTCAGTCGCTCTGTTTATCGTATATGGAAGTAAAAAAGAGATGGTGGCAGA 2161
 QY 342 LeuLysThrLeuGlyIleGlnValThrLysValTyrGluGluGluTyrSerTyrIlePro 361
 Db 2162 TTAGGCGATTTGGGGATCAAAGTGAAGTGTATTAGGAGAGAAATGTGTATGATCCCT 2221
 QY 362 ValGlyGlySerLeuProAsnThrGluGlnLysAsnLeuAlaPheGlyAlaAlaAser 381
 Db 2222 ATGGGAGGACCATCTCCGCGGATTCCTCAAAATGTTATGGCTATGTGGGAATCAGG 2281
 QY 382 MetValHisProAlaThrGlyTyrSerValValArgSerLeuSerGluAlaProLysTyr 401
 Db 2282 ATAGTTTCATCCATCAACAGGTACATGGTGGCTAGGAGCATGGCTTTACCCACGTACTA 2341
 QY 402 AlaSerValIleAlaLysIleLeuLysGlnAspAsnSerAlaTyrValValSerGlyGln 421
 Db 2342 GCTGAAGCCATCGTCGAGGGGCTT-----GGCTCAACAAGAATGATAGAGGGCT 2392
 QY 422 SerSerAlaValAsnIleSerMetGlnAlaTrpSerSerLeuTrpProLysGluArgLys 441
 Db 2393 -----CAACTTACCATAGATTTGGAATGGTTTGGCCTTGGATAGAGA 2440
 QY 442 ArgGlnArgAlaPhePheLeuPheGlyLeuGluLeuIleValGlnLeuAspIleGluAla 461
 Db 2441 TGTGTTAGAGAATGTTATTCATTGGGATGGAGACATGTTGAAGCTGATTTGAAAGG 2500
 QY 462 ThrArgThrPhePheArgThrPhePheArgLeuProThrTrpMetTrpTrpGlyPheLeu 481
 Db 2501 ACTAGGAGATGTTTGACGCTTCTTTGATCTGATCTCAATTAATCTGGCAAGGGTCCCT 2560
 QY 482 GlySerSerLeuSerSerPheAspLeuValLeuPheSerMetTyrMetPheValLeuAla 501
 Db 2561 TCTTCAAGATTTGCTGCAAGAAGCTTGGTTTACTTCAGCTGTGTCTTTTCGCACATGCC 2620
 QY 502 ProAsnSerMetArgMetSerLeuVal 510
 Db 2621 TCAACATCAGCTAGGTGGATATGTT 2647

RESULT 14
 US-09-134-607A-14
 ; Sequence 14, Application US/09134607A
 ; Patent No. 6252141
 ; GENERAL INFORMATION:
 ; APPLICANT: Joseph Hirschberg et al.
 ; TITLE OF INVENTION: POLYNUCLEOTIDES CONTROLLING THE EXPRESSION
 ; OF AND CODING FOR GENE B IN TOMATO AND USE
 ; TITLE OF INVENTION: OF SAME FOR ALTERING CAROTENOID
 ; TITLE OF INVENTION: BIOSYNTHESIS
 ; NUMBER OF SEQUENCES: 25
 ; CORRESPONDENCE ADDRESS:

ADDRESSEE: Mark M. Friedman c/o Anthony Castorina

STREET: 20001 Jefferson Davis Highway, Suite 207
CITY: Arlington
STATE: Virginia
COUNTRY: United States of America
ZIP: 22202

COMPUTER READABLE FORM:
MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk
COMPUTER: Twinhead, Slimnote 890TX
OPERATING SYSTEM: MS DOS version 6.2,
OPERATING SYSTEM: Windows version 3.11
SOFTWARE: Word for Windows version 2.0,
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/134,607A
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Friedman, Mark M.
REGISTRATION NUMBER: 33,883
REFERENCE/DOCKET NUMBER: 325/12
TELECOMMUNICATION INFORMATION:
TELEPHONE: 972-3-562553
TELEFAX: 972-3-562554
TELEX:
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 2876
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

US-09-134-607A-14

Alignment Scores:			
Pred. No.:	Score:	Length:	2876
Best Similarity:	719.50	Matches:	164
Best Local Similarity:	59.17%	Conservative:	78
Query Match:	40.10%	Mismatches:	148
DB:	26.23%	Indels:	19
	4	Gaps:	9
US-09-701-395A-23 (1-529) x US-09-134-607A-14 (1-2876)			
QY	109	AspLeuValValIleGlyCysGlyProAlaGlyLeuSerLeuAlaAlaGluAlaAlaLys	128
Db	1457	GAGGTGATCATTTACGGAGCTGGCCCTGCTGGCTCAGGCTAGCTGACAAAGTTTCTAAA	1516
QY	129	LeuGlyLeuLysValGlyLeuIleGlyProAspLeuProPheThr-----AsnAsp	145
Db	1517	TATGTATTAAAGTGATGTGTGTGACCCCTCA---CCACTCTCCATGTGCCCAAAATAAT	1573
QY	146	TyrGlyValTTPGluAspGluPheLysAspLeuGlyLeuGluATGcysIleGluHISAla	165
Db	1574	TATGTGTTGGGTGATGAGTTTGAGATTAGGACTGGAAATTTGTTAGATCAATAA	1633
QY	166	TrpLysAspThrIleValTyrLeuAspAsnAspAlaProValLeuIleGlyArgAlaTyr	185
Db	1634	TGGCCTATGACTGTGTGCATATAATGATACAAAATAAGTATTTTGGGAAGCCATAT	1693
QY	186	GlyArgValSerArgHisLeuLeuHisGluGluLeuLeuLysArgCysValGluSerGly	205
Db	1694	GGTAGACTTAGTACAAAGAGCTGAAGTTGAAATTTGTAATGTGTGTGAGAACAGA	1753
QY	206	ValSerTyrLeuAspSerLysValGluArgIleThrGluAlaGlyAspGlyHisSerLeu	225
Db	1754	GTAAAGTTTATAAAGCTTAAGTTTGGAAAGTG---GAACATGAAGAATTTGAGTCTTCA	1810
QY	226	ValValCysGluAsnGluIlePheIleProCysArgLeuAlaThrValAlaSerGlyAla	245
Db	1811	ATTGTTTGATGATGGTGAAGATAGAAGGTAGTTTGGTTGGATGCAAGTGGTTTT	1870
QY	246	AlaSerGlyLysLeuLeuGluTyrGluValGlyGlyProArg---ValCysValGlnThr	264

```
Db 1871 GCTAGT---GATTTTATAGATGAC-----AGGCCAAGAAACCATGCTTATCAAT 1921
QY 265 AlaTyrGlyValGluValGluAsnAspProTyrAspProAsnLeuMetValPhe 284
Db 1922 GCTCATGGGTTTGTAGTAAGTTGATATCATCCATTTGATTTGGATAAATGGTGCTT 1981
QY 285 MetAspTyrArgAspTyr-----MetGlnGlnLysLeuGlnCysSerGluGluGlu 301
Db 1982 ATGGATTGAGGATTCCTCATTTGGGTAAATGAGCCATATTTAAGGGTGAATATGCTAAA 2041
QY 302 TyrProThrPheLeuTyrValMetProMetSerProThrArgLeuPhePheGluGluThr 321
Db 2042 GAACCAACATCTTGTATGATCCATTTGATAGAGATTTGTTTCTTGGAGAGACT 2101
QY 322 CysLeuAlaSerLysAspAlaMetProPheAspLeuLysArgLysLeuMetSerArg 341
Db 2102 TCTTGGTGAAGCTGCTCTTTTATCGTATATGGAATGAAGAGAGAGTGGTGGCAAGA 2161
QY 342 LeuLysThrLeuGlyIleGlnValThrLysValTyrGluGluTrpSerTyrIlePro 361
Db 2162 TTAAGGCATTTGGGATCAAGTGAAGTGAAGTGTATTAGGAGACAGAAATGTGTGATCCCT 2221
QY 362 ValGlyGlySerLeuProAsnThrGluGlnLysAsnLeuAlaPheGlyAlaAlaSer 381
Db 2222 ATGGGAGGACCACTTCCGCGGATTCCTCAAAATGTTATGGCTATTGGTGGGAATTCAGGG 2281
QY 382 MetValHisProAlaThrGlyTyrSerValValArgSerLeuSerGluAlaProLysTyr 401
Db 2282 ATAGTTCATCCATCAACAGGATACATGGTGGTAGGAGCATGCTTTAGCACCACTACTA 2341
QY 402 AlaSerValIleAlaLysIleLeuLysGlnAspAsnSerAlaTyrValValSerGlyGln 421
Db 2342 GCTGAAGCATCGTTCGAGGGCTT-----GGCTCAACAAGATGATAAGAGGGTCT 2392
QY 422 SerSerAlaValAsnIleSerMetGlnAlaThrPheSerSerLeuTrpProLysGluArgLys 441
Db 2393 -----CAACTTTACCATAGATTTGGAATGGTTTGGCTTTGGATAGAGA 2440
QY 442 ArgGlnArgAlaPhePheLeuPheGlyLeuGluLeuIleValGlnLeuAspIleGluAla 461
Db 2441 TGTGTAGAGATGTTATTCATTTGGGATGGAGACATGTTGTAAGCTGTATTTGAAGGG 2500
QY 462 ThrArgThrPhePheArgThrPhePheArgLeuProThrTrpMetTrpTrpGlyPheLeu 481
Db 2501 ACTAGAGATTTGTTGACGCTTCTTTGATCTTGATCTTAATCTGCAAGTGGCAAGGTTCTT 2560
QY 482 GlySerSerLeuSerSerPheAspLeuValLeuPheSerMetTrpMetPheValLeuAla 501
Db 2561 TCTTCAAGATTGCTGTCAAGAACTTGTGTTACTCAGCTGTGTCTTTTCGGACATGGC 2620
QY 502 ProAsnSerMetArgMetSerLeuVal 510
Db 2621 TCAACATGACTAGTGGTGAATATGTT 2647
```

RESULT 15

US-08-702-598-1

Sequence 1, Application US/08702598

Patent No. 5850332

GENERAL INFORMATION:

APPLICANT: Camara, Bilal

APPLICANT: Kuntz, Marcel

TITLE OF INVENTION: DNA Constructs, Cells and Plants Derived

TITLE OF INVENTION: Therefrom

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESS: Spencer & Frank

STREET: 1100 New York Avenue, N.W. Suite 300E

CITY: Washington, D.C.

COUNTRY: USA

ZIP: 20005

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

```
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08702,598
APPLICATION NUMBER: US/08702,598
FILING DATE:
CLASSIFICATION: 800
CLASSIFICATION: C12N 15/82
CLASSIFICATION: C12N 9/00
CLASSIFICATION: A01H 5/00
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9403943.5
FILING DATE: 01-MAR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 94400636.1
FILING DATE: 23-MAR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Calvetti, Frederick F.
REGISTRATION NUMBER: 28,557
REFERENCE/DOCKET NUMBER: GROFO 7000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 414-4000
TELEFAX: (202) 414-4040
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1756 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 67..1560
US-08-702-598-1
```

Alignment Scores:

Pred. No.:	8.94e-73	Length:	1756
Score:	705.50	Matches:	158
Percent Similarity:	57.95%	Conservative:	79
Best Local Similarity:	38.63%	Mismatches:	153
Query Match:	25.72%	Indels:	19
DB:	2	Gaps:	9

US-09-701-395a-23 (1-529) x US-08-702-598-1 (1-1756)

```
QY 109 AspLeuValValIleGlyCysGlyProAlaGlyLeuSerLeuAlaAlaGluAlaLys 128
Db 313 GACGTGATCATCATTTGGAACCTGGCCCTCGCGGCTTCGGCTAGCTGAACAGTTCTTAA 372
QY 129 LeuGlyLeuLysValGlyLeuIleGlyProAspLeuProPheThr-----AsnAsn 145
Db 373 TATGGTATTAAAGTATGTTGGCTTGACCTTCA---CCACTTTCCTATGTGGCCAATAT 429
QY 146 TyrGlyValTrpGluAspGluPheLysAspLeuGlyLeuArgCysIleGluHisAla 165
Db 430 TATGGTGTGGTGTGATGACTTTGAAAGTTGGATTAGAAGATTTGCTAGATCATTAAG 489
QY 166 TrpLysAspThrIleValTyrLeuAspAsnAlaProValLeuIleGlyArgAlaTyr 185
Db 490 TGGCTCTGAGTGTGTTTCATATAAGTACCAAGACTAAGTATTTGGACAGACCATAT 549
QY 186 GlyArgValSerArgHisLeuLeuHisLeuLysArgCysValGlySerGly 205
Db 550 GGTAGATTAAGTAAAGAAAGTGAAGTTGAATAGTGTGTTGAAATAAG 609
QY 206 ValSerTyrLeuAspSerLysValGluArgIleThrGluAlaGlyAspGlyHisSerLeu 225
Db 610 GTGAAGTTTATAAGCCCAAGCTTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 666
QY 226 ValValCysGluAsnGluIlePheIleProCysArgLeuAlaThrValAlaSerGlyAla 245
```

Search completed: May 22, 2003, 00:17:48
Job time : 107 secs

GenCore version 5.1.4_p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 21, 2003, 22:12:08 ; Search time 359 Seconds
(without alignments)
3318.404 Million cell updates/sec

Title: US-09-701-395A-23

Perfect score: 2743

Sequence: 1 MELLGVRNLSSCPVWTFGT.....VRHLLSDPSGAVMVRAYLER 529

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-O=Cgn2.1/USPTO.spool/US0701395/runat.15052003.165749.3324/app_query.fasta.1.711
-DB=N_Genseq_101002 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOFCI=0
-LIST=45 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-MODE=LOCAL -OUTFMT=ptc -THR_SCORE=ptc -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-USER=US0701395 -ECGN=1.1.263 -runat.15052003.165749.3324 -NCPU=3 -ICPU=3
-NO_XLPXY -NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Genseq_101002.*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
- 24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2743	100.0	1898	21	AAZ45585	cDNA encoding an A
2	2728	99.5	1848	21	AAZ45595	cDNA encoding an A
3	2728	99.5	1848	21	AAZ46329	Adonis palaestina
4	1959.5	71.4	1860	18	AAZ95371	Arabidopsis thalia
5	1959.5	71.4	1860	21	AAZ45578	cDNA encoding an A
6	1934	70.5	1887	21	AAZ07584	Marigold epsilon-c
7	1931	70.4	1780	21	AAZ46328	Marigold lettuce ep
8	1892	69.0	2426	22	AAZ17465	Spinach lycopene e
9	1518.5	55.4	1377	21	AAZ45586	cDNA encoding a po
10	1394	50.8	1272	19	AAZ19124	Nucleotide sequenc
11	1394	50.8	1272	24	AAZ97356	Brassica napus eps
12	1202	43.8	960	19	AAZ19123	Nucleotide sequenc
13	1202	43.8	962	20	AAZ25064	Brassica napus lyc
14	1202	43.8	962	24	AAZ97355	Brassica napus eps
15	1187	43.3	1039	21	AAZ44887	Arabidopsis thalia
16	897	32.7	720	20	AAZ25065	Brassica napus lyc
17	761	27.7	1942	18	AAZ42979	Capsicum annuum ly
18	750.5	27.4	1650	17	AAZ40207	Tomato lycopene cy
19	750	27.3	1959	21	AAZ07582	Marigold beta-cycl
20	748	27.3	1986	21	AAZ51297	Arabidopsis thalia
21	739	26.9	1590	19	AAZ19125	Nucleotide sequenc
22	739	26.9	1590	20	AAZ25066	Brassica napus lyc
23	739	26.9	1590	24	AAZ97357	Brassica napus bet
24	737	26.9	1614	17	AAZ40206	Tobacco lycopene c
25	733	26.7	1988	21	AAZ37558	Arabidopsis thalia
26	722.5	26.3	2897	21	AAZ51519	L. pennellii lycop
27	722.5	26.3	3265	21	AAZ51521	L. pennellii lycop
28	719.5	26.2	1666	21	AAZ51516	L. esculentum lyco
29	719.5	26.2	1739	21	AAZ51518	L. pennellii lycop
30	719.5	26.2	2876	21	AAZ51517	L. esculentum lyco
31	705.5	25.7	1756	16	AAZ03851	Capsanthin-capsoru
32	703.5	25.6	1740	21	AAZ51520	L. esculentum ogc
33	637.5	23.2	1731	21	AAZ34973	Soybean lycopene c
34	606	22.1	4928	16	AAZ94735	Lycopene cyclase g
35	606	22.1	4928	17	AAZ40205	Cyanobacterial lyc
36	562.5	20.5	1358	21	AAZ34972	Corn lycopene cycl
37	549.5	20.0	563	22	AAZ44247	Physcomitrella pat
38	325	11.8	853	21	AAZ34974	Wheat lycopene cyc
39	155.5	5.7	21091	21	AAZ81523	N. meningitidis pa
40	155.5	5.7	34980	21	AAZ21544	Neisseria meningit
41	155.5	5.7	34980	12	AAZ13722	Neisseria meningit
42	137.5	5.0	1234	17	AAZ13722	Lycopene cyclase -
43	137.5	5.0	1235	17	AAZ40795	Lycopene cyclase c
44	137.5	5.0	1235	18	AAZ91547	Erwinia herbicola
45	137.5	5.0	1893	18	AAZ85207	Squalene epoxidase

ALIGNMENTS

RESULT 1
AAZ45585
ID AAZ45585 standard; cDNA; 1898 BP.

XX AC AAZ45585;

XX DT 06-APR-2000 (first entry)

XX DE cDNA encoding an Adonis palaestina lycopene epsilon-cyclase.

XX KW Lycopene epsilon-cyclase; carotenoid biosynthesis pathway;

XX KW isopentenyl pyrophosphate isomerase; beta-carotene hydroxylase;

XX KW dimethylallyl pyrophosphate; transgenic plant; carotenoid content;

XX KW lethal photooxidative damage; colouring; ss.

XX OS Adonis palaestina.

XX FH Location/Qualifiers

FT CDS 113..1702
 FT /*tag= a
 XX /product= "lycopene epsilon-cyclase"
 PD WO9963055-A1.
 XX 09-DEC-1999.
 XX 02-JUN-1999; 99WO-US12121.
 XX 02-JUN-1998; 98US-0088724.
 PR 02-JUN-1998; 98US-0088725.
 XX (UYMA-) UNIV MARYLAND BALTIMORE.
 PA Cunningham FX, Sun Z;
 XX WPI: 2000-116366/10.
 DR P-PSDB; AAY54290.
 XX New carotenoid biosynthesis genes used to provide transgenic plants
 with altered carotenoid content
 XX Disclosure: Fig 14; 134pp; English.
 CC The present sequence encodes a lycopene epsilon-cyclase protein.
 CC The enzyme is involved in the carotenoid biosynthesis pathway. The
 CC specification also describes other enzymes involved in this pathway,
 CC such as epsilon-cyclase, isopentenyl pyrophosphate (IPP) isomerase and
 CC beta-carotene hydroxylase. IPP isomerase catalyses the reversible
 CC conversion of IPP to dimethylallyl pyrophosphate (DMAPP). Beta-carotene
 CC hydroxylase is responsible for hydroxylating the beta-endgroup in
 CC carotenoids. Lycopene epsilon-cyclase is responsible for the formation
 CC of epsilon-endgroups in carotenoids. The polynucleotides of the
 CC invention are used to provide transgenic plants with altered
 CC carotenoid content. These plants may, for example, may be protected
 CC from lethal photooxidative damage or have altered colouring.
 XX Sequence 1998 BP; 523 A; 342 C; 436 G; 597 T; 0 other;
 SQ

Alignment Scores:
 Pred. No.: 7,35e-282 Length: 1898
 Score: 2743.00 Matches: 529
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 21 Gaps: 0

US-09-701-395A-23 (1-529) x AA245585 (1-1898)

QY 1 MetGluLeuGlyValArgAsnLeuLeuSerSerCysProValTrpThrPheGlyThr 20
 DB 113 ATGGAACCTAGTGTGTCGACACCTATCTCTTCTGCTGTGGACCTTTGGACCA 172
 QY 21 ArgAsnLeuSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 173
 DB 173 AGAAACCTAGTGTGTCGACACCTATCTCTTCTGCTGTGGACCTTTGGACCA 40
 QY 41 ValAspPheGlnValArgAlaAspGlyGlySerGlySerGlySerGlySerGlySerGly 232
 DB 233 GTAGATTTTCAAGTGTGTCGACACCTATCTCTTCTGCTGTGGACCTTTGGACCA 60
 QY 61 GluGlyPheValAspGluGluAspPheLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 292
 DB 293 GAGGCTTTTGGATGAGGAGATTTTATCAAAAGCTGTGTCTGCTGTGGACCTTTGGACCA 80
 QY 81 GlnMetGlnGlnThrLysSerMetGluLysGlnAlaLysLeuAlaAspLysLeuPro 352
 DB 353 CAATGTCAGCAACAAAGTCTATGGAGAACAGCCAGCCAGCCAGCCAGCCAGCCAGCCAG 100
 QY 101 IleProPheGlyGluSerValMetAspLeuValLeuLeuLeuLeuLeuLeuLeuLeu 412
 DB 413 ATACCTTTTGGACCACTGTCGACCTTTGTTGTAATAGTGTGTGGACCTTTGGACCTTT 120
 XX

QY 121 SerLeuAlaAlaGluAlaAlaLysLeuGlyLeuLysValGlyLeuLeuGlyProAspLeu 140
 DB 473 TCACTGGCTGCAGAAAGCTGCTAAGTAGGTTGAAAGTTGGCTTATTGGTCTCTGATCT 532
 QY 141 ProPheThrAsnAsnTyTrpGlyValTrpGluAspGluPheLysAspLeuGlyLeuGly 160
 DB 533 CCTTTTACAATAATATTGTTGGTGGAGACGAGTTCAAAGATCTTGGACTTGAACGT 592
 QY 161 CysIleGluHisAlaTrpLysAspThrIleValTrpLeuAspAsnAspAlaProValLeu 180
 DB 593 TGTATCGACATGCTTGGAAAGCACACCATCTGTATATCTTTGATAATGATGCTCTGCT 652
 QY 181 IleGlyArgAlaTyTrpGlyArgValSerArgHisLeuLeuHisGluGluLeuLysArg 200
 DB 653 ATTGGTGTGCATATGAGGAGTAGTGTGACATTTGTACATGAGGAGTTGCTGAAAGG 712
 QY 201 CysValGluSerGlyValSerTyTrpLeuAspSerLysValGluArgIleThrGluAlaGly 220
 DB 713 TGTGTGGAGTCAGGTGTATCATATCTTGTATCTTAAAGTGAAAGGATCACTGAAGCTGT 772
 QY 221 AspGlyHisSerLeuValValCysGluAsnGluIlePheProCysArgLeuAlaThr 240
 DB 773 GATGGCCATAGCTTGTAGTTTGTGAAATGAGATCTTTATCCCTTGCAGGCTTCTACT 832
 QY 241 ValAlaSerGlyAlaAlaSerGlyLysLeuGluTyTrpGluValGlyGlyProArgVal 260
 DB 833 GTTGCATCTGGACGACCTTCAGGAAACTTTTGGAGTATGAGTAGGTGGCCCTCGTGT 892
 QY 261 CysValGlnThrAlaTyTrpGlyValGluValGluValGluAsnAsnProTyTrpAsp 280
 DB 893 TGTGTCCAACGCTTATGGGTGTGAGTGTGAGTGTGAGGAGAACTATCATGATCCCAAC 952
 QY 281 LeuMetValPheMetAspTyTrpArgAspTyTrpMetGlnGlnLysLeuGlnCysSerGlu 300
 DB 953 TTAATGTGTATCATGAGTGTATGAGTGTATGAGTGTATGAGTGTATGAGTGTATGAG 1012
 QY 301 GluTyTrpProThrPheLeuTyTrpValMetProMetSerProThrArgLeuPheGlyGlu 320
 DB 1013 GAATATCCAACTTCTTCTATGTCTATGTCATGCCCATCTGCCCAACAGACTTTTTCAG 1072
 QY 321 ThrCysLeuAlaSerLysAspAlaMetProPheAspLeuLysLeuLysLeuMetSer 340
 DB 1073 ACCTGTTGGCTTCAAAGATGCCATGCCATCTGATCTACTCAAGAGAAACTGATGTCA 1132
 QY 341 ArgLeuLysThrLeuGlyIleGlnValThrLysValTyTrpGluGluTrpSerTyTrp 360
 DB 1133 CGATTGACACTCTGGGTATCCAAAGTTTACAAAGTTTATGAGGAGATGATGATATATT 1192
 QY 361 ProValGlyGlySerLeuProAsnThrGluGlnLysAsnLeuAlaPheGlyAlaAla 380
 DB 1193 CCTGTTGGTGTCTTCTTACCAACACAGAGCAAAAGAACCTTAGCATTTGGTGTGACGA 1252
 QY 381 SerMetValHisProAlaThrGlyTyTrpSerValValArgSerLeuSerGluAlaProLys 400
 DB 1253 AGCATGTGTGCATCCAGCAACAGGCTATTCGTTGTGACGGTCACTGTCCAGAACCTCCAAA 1312
 QY 401 TyrAlaSerValIleAlaLysIleLeuLysGlnAspAsnSerAlaTyTrpValSerGly 420
 DB 1313 TATCCTCTCTGTAATGCAAGATTTTGAAGCAAGATTAACCTCGGTATGTTGTTCTGGA 1372
 QY 421 GlnSerSerAlaValAsnIleSerMetGlnAlaTrpSerSerLeuTrpProLysGluArg 440
 DB 1373 CAAAGTAGTCAGTAAACATTTCAATGCAAGCATGAGCAGTCTTTGGCCAAAGGAGCGA 1432
 QY 441 LysArgGlnArgAlaPhePheLeuPheGlyLeuGluLeuIleValGlnLeuAspIleGlu 460
 DB 1433 AAACGTCAAGAGCATCTTCTTTTGGATTAGAGCTTATTGTGACGCTAGATATTGAA 1492
 QY 461 AlaThrArgThrPhePheArgThrPhePheArgLeuProThrTrpMetTrpTrpGlyPhe 480
 DB 1493 GCAACCAAGAACATCTTTAGAACCTTCTTCCGCTTGCACACTTGTGATGTGTTGGGTTC 1552

QY 481 LeuGlySerSerLeuSerSerPheAspLeuValLeuPheSerMetTyrMetPheValLeu 500
 Db 1553 CTTGGGCTTCACATATCATCTTTTCGATCTGCTCTGTTTTCATGATCATGTTGTTTG 1612
 QY 501 AlaProAsnSerMetArgMetSerLeuValArgHisLeuLeuSerAspProSerGlyAla 520
 Db 1613 GCCCCAACAGCAGAGATGTCACCTTGTGAGACATTTGCTTTCAGATCCCTTCGGTGCA 1672
 QY 521 ValMetValArgAlaTyrLeuGluArg 529
 Db 1673 GTTATGCTAAGAGCTTACCTCGAAAGG 1699

RESULT 2

AAZ45595
 ID AAZ45595 standard; cDNA; 1848 BP.

XX
 AC AAZ45595;

DT 06-APR-2000 (first entry)

XX cDNA encoding an Adonis palaeatina lycopene epsilon-cyclase.

XX Lycopene epsilon-cyclase; carotenoid biosynthesis pathway;
 KW isopentenyl pyrophosphate isomerase; beta-carotene hydroxylase;
 KW dimethylallyl pyrophosphate; transgenic plant; carotenoid content;
 KW lethal photooxidative damage; colouring; ss.

XX Adonis palaeatina.

XX Key Location/Qualifiers
 FT CDS 116..1705

FT /*tag= a

FT /product= "lycopene epsilon-cyclase"

FT /transl_except= (pos: 278..280, aa: Ser)

FT /transl_except= (pos: 743..745, aa: Asp)

FT /transl_except= (pos: 806..808, aa: Glu)

FT /transl_except= (pos: 1166..1168, aa: Val)

FT /transl_except= (pos: 1682..1684, aa: Arg)

XX W09963055-A1.

XX PN

XX PD

XX 09-DEC-1999.

XX XX

PF 02-JUN-1999; 99WO-US12121.

XX XX

PR 02-JUN-1998; 98US-0088724.

PR 02-JUN-1998; 98US-0088725.

XX XX

PA (UYMA-) UNIV MARYLAND BALTIMORE.

XX XX

PI Cunningham FX, Sun Z;

XX WPI; 2000-116366/10.

DR P-PSDB; AAY54303.

XX XX

PT New carotenoid biosynthesis genes used to provide transgenic plants
 with altered carotenoid content

XX XX

PS Disclosure; Fig 23; 134pp; English.

XX XX

CC The present sequence encodes a lycopene epsilon-cyclase protein.
 CC The enzyme is involved in the carotenoid biosynthesis pathway. The
 CC specification also describes other enzymes involved in this pathway,
 CC such as epsilon-cyclase, isopentenyl pyrophosphate (IPP) isomerase and
 CC beta-carotene hydroxylase. IPP isomerase catalyses the reversible
 CC conversion of IPP to dimethylallyl pyrophosphate (DMAPP). Beta-carotene
 CC hydroxylase is responsible for hydroxylating the beta-endgroup in
 CC carotenoids. Lycopene epsilon-cyclase is responsible for the formation
 CC of epsilon-endgroups in carotenoids. The polynucleotides of the
 CC invention are used to provide transgenic plants with altered
 CC carotenoid content. These plants may, for example, may be protected
 CC from lethal photooxidative damage or have altered colouring.

XX

SQ Sequence 1848 BP; 522 A; 334 C; 420 G; 572 T; 0 other;

Alignment Scores:

Pred No.: 2 8e-280 Length: 1848
 Score: 2728.00 Matches: 524
 Percent Similarity: 100.00% Conservative: 5
 Best Local Similarity: 99.05% Mismatches: 0
 Query Match: 99.45% Indels: 0
 DB: 21 Gaps: 0

US-09-701-395a-23 (1-529) x AAZ45595 (1-1848)

QY 1 MetGluLeuLeuGlyValArgAsnLeuLeuSerSerCysProValTrpThrPheGlyThr 20

Db 116 ATGGAACCTACTGGTGTTCGCAACCTCATCTCTCTGCCCCTGCTGGACCTTTTGAACA 175

QY 21 ArgAsnLeuSerSerSerLysLeuAlaTyrAsnIleHisArgTyrClySerSerCysArg 40

Db 176 AGAAACCTTAGTAGTTCAAAACCTAGCTTATAACATACATCATGATATGGTTCCTTCTGTAGA 235

QY 41 ValAspPheGlnValArgAlaAspGlyGlySerGlySerArgSerSerValAlaTyrLys 60

Db 236 GTAGATTTTCAAGTCAGGCGCTGATGGTGAACGGGAGTAGAACCTCTCTGCTTATATAA 295

QY 61 GluGlyPheValAspGluAspPheIleLysAlaGlyGlySerGluLeuLeuPheVal 80

Db 296 GAGGTTTGTGGACGAGGAGGATTTATCAAAAGCTGGTGTCTGAGCTTTTGTGTGTC 355

QY 81 GlnMetGlnGlnThrLysSerMetGluLysGlnAlaLysLeuAlaAspLysLeuProPro 100

Db 356 CAAATGCGAGCAACCAAGCTCTATGGAGAAACAGGGCAAGCTCGCGATTAAGTTGCCACCA 415

QY 101 IleProPheGlyGluSerValMetAspLeuValIleGlyCysGlyProAlaGlyLeu 120

Db 416 ATACCTTCGGAGATCTGTGATGCATCTGGTGTGTAATAGGTGTGGACCTGCTGCTCTT 475

QY 121 SerLeuAlaAlaGluAlaLysLeuGlyLeuLysValGlyLeuIleGlyProAspLeu 140

Db 476 TCACTGGCTGCAGAAAGCTCTAAGCTAGGCTTGAAGTTGGGCTTATTTGCTGCTGCTT 535

QY 141 PropheThrAsnAsnTyrGlyValTrpGluAspGluPheLysAspLeuGlyLeuArg 160

Db 536 CCTTTTACAATTAATATGGTGTGGGGAAGACGAGTTTCAAAGATCTTGGACTTGAACGT 595

QY 161 CysIleGluHisAlaTrpLysAspThrIleValTyrLeuAspAsnAspAlaProValLeu 180

Db 596 TGTATCGAGCATGCTTGGAAAGGACACCATCGTATATCTTTGACAATGATGCTCTCTGCTT 655

QY 181 IleGlyArgAlaTyrGlyArgValSerArgHisLeuLeuHisGluLeuLeuLysArg 200

Db 656 ATTGCTCGTCATATGAGAGGATTTAGCCGGCATTTGCTGCATGAAGAGTTGCTGAAAGG 715

QY 201 CysValGluSerGlyValSerTyrLeuAspSerLysValGluArgIleThrGluAlaGly 220

Db 716 TGTGTCAGTTCAGGTGTATCATATCTGAATCTTAAAGTGAAGGATCATCTGAGGCTGGT 775

QY 221 AspGlyHisSerLeuValValCysGluAsnGluIlePheIleProCysArgLeuAlaThr 240

Db 776 GATGCCATAGTCTGTAGTTGTGAAAACGACATCTTATCCCTTGCAGGCTGCTACT 835

QY 241 ValAlaSerGlyAlaAlaSerGlyLysLeuLeuGluTyrGluValGlyGlyProArgVal 260

Db 836 GTTGTCATCTGGAGCAGCTTCAGGGAACCTTTTGGAGTAGTGAAGTAGGTGGCCCTCGTGT 895

QY 261 CysValGlnThrAlaTyrGlyValGluValGluAlaAsnAsnProTyrAspProAsn 280

Db 896 TGTGTCCAACTGCTTATGGTGTGAGGTTGAGGTGGAGAACAAATCCATACCATCCCCAAC 955

QY 281 LeuMetValPheMetAspTyrArgAspTyrMetGlnGlnLysLeuGlnCysSerGluGlu 300

Db 956 TTATGGTATTTATGGACTACAGACATATATGCAACAGAAATTTACAGTGTCTGGAGAA 1015

QY 301 GluTyrProThrPheLeuTyrValMetProMetSerProThrArgLeuPhePheGluGlu 320

QY 101 IleProPheGlyCysLeuValMetAspLeuValValIleGlyCysGlyProAlaGlyLeu 120
DB 416 ATACCTTTCCGAGAACTCTGTGGAGCTTGGTTGTAATAGGTGTGGACCTGTGCTCT 475
QY 121 SerLeuAlaAlaGluAlaAlaLysLeuGlyLeuLysValGlyLeuIleGlyProAspLeu 140
DB 476 TCACCTGGCTGCAAGAGCTCTAGCTAGGCTTGAAGTTGGCTTATGGTCTGATCTT 535
QY 141 ProPheThrAsnAsnTyrGlyValTyrGluAspGluPheLysAspLeuGlyLeuGluArg 160
DB 536 CCTTTTACAAATATATATGGTGTGGAGAGAGCTTCAAGATCTTGGACTTGAAGCT 595
QY 161 CysIleGluHisAlaTrpLysAspThrIleValTyrLeuAspAsnAlaProValLeu 180
DB 596 TGTATCGAGCTGCTTGGAGGACACCATCGTATATCTTACAAATGATGCTCTGCTCT 655
QY 181 IleGlyArgAlaTyrGlyArgValSerArgHisLeuLeuHisGluLeuLysArg 200
DB 656 ATGGTGTGTCATATGGAGAGTATGCGGCACTTCTGCATGAAGAGTTGCTGAAAGG 715
QY 201 CysValGluSerGlyValSerTyrLeuAspSerLysValGluArgIleThrGluAlaGly 220
DB 716 TGTGTGAGCTAGGTATCATATCTGAATCTTAAAGTGAAGAGTCACTGAAGCTGGT 775
QY 221 AspGlyHisSerLeuValValCysGluAsnGluIlePheIleProCysArgLeuAlaThr 240
DB 776 GATGGCATAGCTCTGTAGCTTGTGAAACGACATCTTATCCCTTGCAGGCTTGCTACT 835
QY 241 ValAlaSerGlyAlaAlaSerGlyLysLeuLeuGluTyrGluValGlyProArgVal 260
DB 836 GTTGCATCTGGAGCAGCTTCAAGGAACTTTTGGAGTATGAAGTAGTGCCCTCGTGT 895
QY 261 CysValGlnThrAlaTyrGlyValGluValGluValGluAsnProTyrAspProAsn 280
DB 896 TGTGTCCAACTGCTTATGTGTGGAGGTTGAGTGGAGAACATCCATACGATCCCAAC 955
QY 281 LeuMetValPheMetAspTyrArgAspTyrMetGlnGlnLysLeuGlnCysSerGluGlu 300
DB 956 TTAATGGTATTTATGACATACAGACTATATGCAACAGAAATTACAGTCTCGAAGAA 1015
QY 301 GluTyrProThrPheLeuTyrValMetProMetSerProThrArgLeuPheGluGlu 320
DB 1016 GAATATCCAACTTCTCTATGTATGATGCCATGTGCCCAACAGACTTTTTTTTGAGGAA 1075
QY 321 ThrCysLeuAlaSerLysAspAlaMetProPheAspLeuLeuLysArgLysLeuMetSer 340
DB 1076 ACCTGTTGGCCCTCAAAAGATGCCATGCTTTCGATCTACTGAGAGAAACTAAATGTCA 1135
QY 341 ArgLeuLysThrLeuGlyIleGlnValThrLysValTyrGluGluTrpSerTyrIle 360
DB 1136 CGATTGAAGACTCTGGGTATCCAAAGTTACAAAAATTTATGAAGAGGAATGGTCTTATAT 1195
QY 361 ProValGlyGlySerLeuProAsnThrGluGlnLysAsnLeuAlaPheGlyAlaAlaAla 380
DB 1196 CTTGTTGGGGGTTCTTTACCAACAGAGCAAAAGAACTAGCATTTGGTGTGAGCA 1255
QY 381 SerMetValHisProAlaThrGlyTyrSerValValArgSerLeuSerGluAlaProLys 400
DB 1256 AGCATGGTGCATCCAGCAAGCTATTCTGGTTGTACGATCACTATCAGAGCTCCAAA 1315
QY 401 TyrAlaSerValIleAlaLysIleLeuLysGlnAspAsnSerAlaTyrValValSerGly 420
DB 1316 TATGCTTCTGTAATTTGCAAGATTTTGAAGCAAGATAAATCTCTGCATATGGTTCTGGA 1375
QY 421 GlnSerSerAlaValAsnIleSerMetGlnAlaTrpSerSerLeuTrpProLysGluArg 440
DB 1376 CAAGCAGTGCAGTAAACATTTCAATGCAAGCATGGAGCAGTCTTTGGCCAAAGAGCGA 1435
QY 441 LysArgGlnArgAlaPhePheLeuPheGlyLeuGluLeuIleValGlnLeuAspIleGlu 460
DB 1436 AAACGTCAAGAGCATCTCTTTTCGGGTTAGACCTTATGTGCGAGTAGATATGAA 1495
QY 461 AlaThrArgThrPhePheArgThrPhePheArgLeuProThrTrpMetTrpTrpGlyPhe 480

DB 1496 GCACCAAGAGCTTCTTTAGAACCTTCTTCGCTTGCCAACTTGCGATGTGGTGGGTTTC 1555
QY 481 LeuGlySerSerLeuSerSerPheAspLeuValLeuPheSerMetTyrMetPheValLeu 500
DB 1556 CTTGGGCTTTCACATATCATCTTTTCGATCTTGTATGTTTTCATGTATCATGTTTGTG 1615
QY 501 AlaProAsnSerMetArgMetSerLeuValArgHisLeuLeuSerAspProSerGlyAla 520
DB 1616 GCCCCGAACAGCATGAGGATGTCATCTGTGAGACATTTGCTTTCAGATCCCTCTGCTGCA 1675
QY 521 ValMetValArgAlaTyrLeuGluArg 529
DB 1676 GTTATGTTAAAGCTTACCTCGAAAGG 1702

RESULT 4

AA95371
ID AA95371 standard; cDNA; 1860 BP.

AC AA95371;

DT 14-APR-1998 (first entry)

XX Arabidopsis thaliana epsilon cyclase cDNA.

DE Epsilon cyclase; E-cyclase; carotenoid; biosynthetic enzyme;
KW pigment; vector; pATeps; ss.

XX Arabidopsis thaliana.

XX Key Location/Qualifiers
FH CDS 109..1683

FT polyA_signal 1732..1736

FT /*tag= a

FT /*tag= b

XX WO9736998-A1.

XX 09-OCT-1997.

XX 28-JAN-1997; 97WO-US00540.

XX 29-MAR-1996; 96US-0624125.

XX (UYMA-) UNIV MARYLAND BALTIMORE.

XX Cunningham FX, Sun Z;

XX WPI; 1997-503091/46.

XX P-PSDB; AAW35346.

XX Eukaryotic carotenoid biosynthetic enzymes and related genes -
useful to control ratio of various carotenoid(s) in host and for
production of novel carotenoid pigments

XX Claim 4; Page 33-35; 89pp; English.

XX This cDNA sequence codes for Arabidopsis thaliana epsilon cyclase
(see AAW35346), an enzyme responsible for the formation of epsilon
end-groups in carotenoids. It was isolated by mass excision and
colour complementation screening of an A. thaliana cDNA library.
Claimed expression vector pATeps comprising the epsilon cyclase
gene is deposited as ATCC 98005. Claimed isolated DNA sequences
(see AA95371-73 and AA95389-91) encoding the claimed eukaryotic
carotenoid biosynthetic enzymes epsilon cyclase, beta-carotene
hydroxylase and isopentenyl pyrophosphate isomerase (see
AAW35346-51) are used in methods for augmenting the accumulation of
carotenoids and for the production of novel and rare carotenoids
in host cells. Methods are also provided for controlling the ratio
of various carotenoids in a host, and for screening for eukaryotic
genes that encode enzymes of carotenoid biosynthesis and metabolism.

XX Sequence 1860 BP; 531 A; 327 C; 448 G; 554 T; 0 other;

Alignment Scores:

Pred. No.: 2,14e-198 Length: 1860
 Score: 1959.50 Matches: 376
 Percent Similarity: 81.65% Conservative: 60
 Best Local Similarity: 70.41% Mismatches: 81
 Query Match: 71.44% Indels: 17
 DB: 18 Gaps: 5

US-09-701-395A-23 (1-529) x AAT95371 (1-1860)

QY	1	MetGluLeuLeuGluValArgAsnLeu-----lSeSerSerCysProValTrp	16
DB	109	ATGGAGTCTGTGGGCTAGGAATTCGCAGCAATGGCGGTTTCAACATTTCCGTATCG	168
QY	17	ThrPheGlyThrArgAsnLeuSerSerLysLeuAlaThrAsnIleHisArgTyrGly	36
DB	169	AGTTGTGAGGAAATTCAGTGGTAAAGAGATACAGCTATAGGAATATTCGTTTCGGT	228
QY	37	SerSerCysArgValAspPheGlnValArgAlaAspGlyGly-----SerGlySerArg	54
DB	229	---TTGTGT-----AGTGTACAGCTAGCGCGCGGAGTTCGGTAGTGAG	273
QY	55	SerSerValAlaTyrLysGluGlyPheValAspGluGluAspPheIleLysAlaGlyCly	74
DB	274	AGTTGTGACGGTGAGAGAAATTCGCTGACGAGGAAGATTTGTGAAAGCTGGTGGT	333
QY	75	SerGluLeuLeuPheValGlnMetGlnGlnThrLysSerMetGluLysGlnAlaLysLeu	94
DB	334	TCGAGATTCATTTGTTCAATGCAGCAGACAAAGATATGGATGAACATCTTAAGCTT	393
QY	95	AlaAspLysLeuProPheGlyPheValMetAspLeuValValIleGly	114
DB	394	GTTGATAAGTTGCCCTTATCAATGGTGTATGCTTGGATCATGTGGTATTTGGT	453
QY	115	CysGlyProAlaGlyLeuSerLeuAlaAlaGluAlaLysLeuGlyLysValGly	134
DB	454	TGTGCTGCTGCTGTGTAGCTTGGCTGCAGATCATGCTTGGATTAAGAGTTGA	513
QY	135	LeuIleGlyProAspLeuProPheThrAsnAspTyrGlyValTrpGluAspGluPheLys	154
DB	514	CTCATGTTGTCAGATCTCTCTTTACTAACAAATTAGCGTGTGGGAAGATCAATCAAT	573
QY	155	AspLeuGlyLeuGluArgCysIleGluHisAlaTrpLysAspThrIleValTyrLeuAsp	174
DB	574	GATCTGGGCTGCAAAAGATGATTAGCATGTTTGGAGAGACATTTGTGTATCTGGAT	633
QY	175	AsnAspAlaProValLeuIleGlyArgAlaTyrGlyArgValSerArgHisLeuHis	194
DB	634	GATGACAAGCTATTACCATTTGCCGCTGCTTATGGAAGAGTATGCTGAGCTTTCCTCAT	693
QY	195	GluGluLeuLeuLysArgCysValGluSerGlyValSerTyrLeuAspSerLysValGlu	214
DB	694	GAGGAGCTTTGAGAGGTTGTGAGTACAGTGTCTGCTACCTTACCTGAAAGTTGAC	753
QY	215	ArgIleThrGluAlaGlyAspGlyHisSerLeuValValCysGluAsnGluIlePheIle	234
DB	754	AGCATAAACAGAGCTTCATGAGGCTTAGACTTGTGTCAGCAGCAATAACGTCAT	813
QY	235	ProCysArgLeuAlaThrValAlaSerGlyAlaAlaSerGlyLysLeuGluTyrGlu	254
DB	814	CCCTGCAGGCTTGCCTGCTGCTGAGCAGCTTCGGAAAGCTCTTGAATACGAA	873
QY	255	ValGlyGlyProArgValCysValGlnThrAlaTyrGlyValGluValGluAsn	274
DB	874	GTTGGTGGAGCTAGAGTCTGTGCAAACTGCATCGGCTGGAGTTGAGGTGCAAAAT	933
QY	275	AsnProTyrAspProAsnLeuMetValPheMetAspTyrArgAspTyrMetGlnGlnLys	294
DB	934	AGTCCATATGATCCAGATCAATGTTTTCATGGATTACAGAGATTATACAAACGAGAA	993
QY	295	LeuGlnCysSerGluGluTyrProThrPheLeuTyrValMetProMetSerProThr	314
	

DB 994 GTTCGGAGCTTAGAAGCTCAGTATCCAACGTTTCTGTACGCCATGCCCTATGACAAAGTCA 1053
 QY 315 ArgLeuPhePheGluGluThrCysLeuAlaSerLysAspAlaMetProPheAspLeuLeu 334
 DB 1054 AGACTCTCTTCGAGGAGACATGTTTGGCTCAAAAGATGTCATGCCCTTTGATTGCTA 1113
 QY 335 LysArgLysLeuMetSerArgLeuLysThrLeuGlyIleGlnValThrLysValTyrGlu 354
 DB 1114 AAAAGAAAGCTCATGTTAAAGATTAGATACCTCGAATTCGAATTTAAAGACTTACGAA 1173
 QY 355 GluGluTrpSerTyrIleProValGlySerLeuProAsnThrGluGlnLysAsnLeu 374
 DB 1174 GAGAGTGGTCTATATCCAGTTGGTGTCTCTGCCAAACACCAAGAAATCTC 1233
 QY 375 AlaPheGlyAlaAlaSerMetValHisProAlaThrGlyTyrSerValValArgSer 394
 DB 1234 GCCTTTGGTCTCGCTAGCATGGTACATCCGCAACAGGCTATTTCAGTTGTGAGACT 1293
 QY 395 LeuSerGluAlaProLysTyrAlaSerValIleAlaLysIleLeuLysGlnAspAsnSer 414
 DB 1294 TTCTCTGAAGCTCCAAATATCATCATCGATCGCAGAGATCTTAAGAGAAGACTACC 1353
 QY 415 AlaTyrValValSerGlyGlnSerSerAlaValAsnIleSerMetGlnAlaTrpSerSer 434
 DB 1354 AACAGATCAACAGT-----AATATTTCAAGACAGCTTGGGATACT 1395
 QY 435 LeuTrpProLysGluArgLysArgGlnArgAlaPhePheLeuPheGlyLeuGluLeu 454
 DB 1396 TTATGCCACCAGAAAGAAAGACAGAGACTTCTTCTTCTTGGTCTTGCATCAT 1455
 QY 455 ValGlnLeuAspIleGluAlaThrArgThrPhePheArgThrPhePheArgLeuProThr 474
 DB 1456 GTTCAATTCGATCCGAGGATTTAGAGTCAACATTAAACATCAGGAGATCTCGTTCTTGT 1515
 QY 475 TrpMetTrpTrpGlyPheLeuGlySerSerLeuSerSerPheAspLeuValLeuPheSer 494
 DB 1516 TGGATGTGGCAAGGTTTCTAGGATCAACATTAAACATCAGGAGATCTCGTTCTTGT 1575
 QY 495 MetTyrMetPheValLeuAlaProAsnSerMetArgMetSerLeuValArgHisLeuLeu 514
 DB 1576 TTATACATGTTCTCATTTTCCAAACAAATTTGAGAAAGGTCATCATCATCTCATC 1635
 QY 515 SerAspProSerGlyAlaValMetValArgAlaTyrLeuGlu 528
 DB 1636 TCTGATCCAAACGGAGCAACCATGATAAAACCTATCTCAA 1677

RESULT 5
 AA245578
 ID AA245578 standard; cDNA; 1860 BP.

XX AA245578;
 XX
 DT 06-APR-2000 (first entry)
 XX cDNA encoding an Arabidopsis thaliana lycopene epsilon-cyclase.
 DE Lycopene epsilon-cyclase; carotenoid biosynthesis pathway.
 KW isopentenyl pyrophosphate isomerase; beta-carotene hydroxylase;
 KW dimethylallyl pyrophosphate; transgenic plant; carotenoid content;
 KW lethal photooxidative damage; colouring; ss.
 OS Arabidopsis thaliana.

Key Location/Qualifiers
 CDS 109..1683
 /tag= a
 /product= "lycopene epsilon-cyclase"

WO9963055-A1.

09-DEC-1999.

XX
 XX 02-JUN-1999; 99WO-US12121.

XX 02-JUN-1998; 98US-0088724.
 PR 02-JUN-1998; 98US-0088725.
 XX (UYMA-) UNIV MARYLAND BALTIMORE.
 PA Cunnigham FX, Sun 2;
 PI WPI; 2000-116366/10.
 DR P-PSDB; AAY54283.
 XX New carotenoid biosynthesis genes used to provide transgenic plants
 PT with altered carotenoid content -
 XX Example; Page 61-63; 134pp; English.
 XX The present sequence encodes a lycopene epsilon-cyclase protein.
 CC The enzyme is involved in the carotenoid biosynthesis pathway. The
 CC specification also describes other enzymes involved in this pathway,
 CC such as epsilon-cyclase, isopentenyl pyrophosphate (IPP) isomerase and
 CC beta-carotene hydroxylase. IPP isomerase catalyses the reversible
 CC conversion of IPP to dimethylallyl pyrophosphate (DMAPP). Beta-carotene
 CC hydroxylase is responsible for hydroxylating the beta-endgroup in
 CC carotenoids. Lycopene epsilon-cyclase is responsible for the formation
 CC of epsilon-endgroups in carotenoids. The polynucleotides of the
 CC invention are used to provide transgenic plants with altered
 CC carotenoid content. These plants may, for example, may be protected
 CC from lethal photooxidative damage or have altered colouring.
 XX
 SQ Sequence 1860 BP; 531 A; 327 C; 448 G; 554 T; 0 other;

Alignment Scores:
 Pred. No.: 2,14e-198 Length: 1860
 Score: 1959.50 Matches: 376
 Percent Similarity: 81.65% Conservative: 60
 Best Local Similarity: 70.41% Mismatches: 81
 Query Match: 71.44% Indels: 17
 DB: 21 Gaps: 5

US-09-701-395a-23 (1-529) x AAZ45578 (1-1860)

QY 1 MetGluLeuGlyValArgAsnLeu-----lleSerSerCysProValTrp 16
 DB 109 ATGGAGTGTGGTGGCGTGAAGATTTCGCAGCAATGCGGTTTCAACATTTCCGTCATGG 168
 QY 17 ThrPheGlyThrArgAsnLeuSerSerSerLysLeuAlaTyrAsnIleHisArgTyrGly 36
 DB 169 AGTTGTCGAAGGAAATTTCCAGTGGTTAAGAGATACAGCTATAGGAATATTCGTTTCGGT 228
 QY 37 SerSerCysArgValaspPheGlnValArgAlaaspGlyGly-----SerGlySerArg 54
 DB 229 ---TTGTGT-----AGTGTCAAGCTAGCGCGCGGGAAGTTCCGGTGTAGTGAG 273
 QY 55 SerSerValAlaTyrLysGluGlyPheValaspGluaspPheIleLysAlaGlyGly 74
 DB 274 AGTTGTAGCGGTGAGAGAGATTTCGTCAGCAAGAAATTTTGTGAAGCTGGTGGT 333
 QY 75 SerGluLeuPheValGlnMetGlnThrLysSerMetGluLysGlnAlaLysLeu 94
 DB 334 TCTGAGATTCTATTGTTCAATGTCAGCAGCAACAAGATATGATGAACAGCTTAAGCTT 393
 QY 95 AlaAspLysLeuProPheProPheGlyGluSerValMetAspLeuValIleGly 114
 DB 394 GTTGATAAGTGTGCTCCTATATCAATTTGGTGTGCTTTGGATCATGTGTTATGGT 453
 QY 115 CysGlyProAlaGlyLeuSerLeuAlaAlaGluAlaLysLeuGlyLeuLysValGly 134
 DB 454 TGTGGTCTCTGCTGCTAGCTGGTGCAGAAATCTTGGTGTGCTTGGTGTGCTTGGT 513
 QY 135 LeuIleGlyProAspLeuProPheThrAsnAsnTyrGlyValTrpGluAspGluPheLys 154
 DB 514 CTATTTGGTCCAGATCTTCTTTACTTAACTTACAAATACCGTGTGTTGGGAGATGAATTCAT 573

QY 155 AspLeuGlyLeuGluArgCysIleGluHisAlaTrpLysAspThrIleValTyrLeuAsp 174
 DB 574 GATCTTGGCTCGCAAAATGTTATGAGCATGTTTGGAGAGAGACTATTGTGTATCTGGAT 633
 QY 175 AsnAspAlaProValLeuIleGlyArgAlaTyrGlyArgValSerArgHisLeuLeuHis 194
 DB 634 GATGACAAGCCTATTACCATTTGCCGCTGCTTATGGAAGAGATTAGTCGACGTTTGCCTCAT 693
 QY 195 GluGluLeuLeuLysArgCysValGluSerGlyValSerGlyValSerGlyValGlu 214
 DB 694 GAGGAGCTTTTGGAGGAGTGTCTCGAGTCAGGTGCTCTGCTACCTTACCTCGAAGTTGAC 753
 QY 215 ArgIleThrGluAlaGlyAspGlyHisSerLeuValValCysGluAsnGluIlePheIle 234
 DB 754 AGCAATAACAGAAAGCTTCTGATGCCCTTAGACTTGTGCTGTGACGACAAATAACGTCATT 813
 QY 235 ProCysArgLeuAlaThrValAlaSerGlyValAlaSerGlyLysLeuGluTyrGlu 254
 DB 814 CCTCGAGCTTGGCAGCTGCTTCTTGGAGCAGCTTCGGGAAGCTCTTGCATTAACGAA 873
 QY 255 ValGlyGlyProArgValCysValGlnThrAlaTyrGlyValGluValGluValGluAsn 274
 DB 874 GTTGTGGACCTAGAGTCTGTCTGCAACTGTCATACGGGTGGAGGTGGGAAAT 933
 QY 275 AsnProTyrAspProAsnLeuMetValPheMetAspTyrArgAspTyrMetGlnLys 294
 DB 934 AGTCCATATGATCCAGATCAAAATGGTTTTCATGGATTACAGAGATTATATAACGAGAAA 993
 QY 295 LeuGlnCysSerGluGluTyrProThrPheLeuTyrValMetProMetSerProThr 314
 DB 994 GTTCGGAGCTTAGAAGCTGAGTATCCACGCTTCTCTACGCCATGCTATGACAAAGTCA 1053
 QY 315 ArgLeuPhePheGluGluThrCysLeuAlaSerLysAspAlaMetProPheAspLeuLeu 334
 DB 1054 AGACTCTTCTCGAGAGACATGTTGGCTCAAAAGATGTCATGCGCTTGTATTGCTA 1113
 QY 335 LysArgLysLeuMetSerArgLeuLysThrLeuGlyIleGlnValThrLysValTyrGlu 354
 DB 1114 AAACGAAGCTCATGTTAAGATTAGATACACTCGGAATTCGAATTTAAAGACTTACGAA 1173
 QY 355 GluGluTrpSerTyrIleProValGlyGlySerLeuProAsnThrGluGlnLysAsnLeu 374
 DB 1174 GAGGAGTGGTCTATATCCAGTTGGTGGTCTCTTCCCAACACCCAGCAAAAGAAATCTC 1233
 QY 375 AlaPheGlyAlaAlaSerMetValHisProAlaThrGlyTyrSerValValArgSer 394
 DB 1234 GCTTTGGTGTGCTGCGCTAGCATGCCGCAACAGGCTATTTCAGTTGTGAGATCT 1293
 QY 395 LeuSerGluAlaProLysTyrAlaSerValIleAlaLysIleLeuLysGlnAspSer 414
 DB 1294 TTGCTGAAGCTCCAAAATATATCATCGATCATCGCAGAGATATCTAAGAGAAGAGACTACC 1353
 QY 415 AlaTyrValValSerGlyGlnSerSerAlaValAsnIleSerMetGlnAlaTrpSerSer 434
 DB 1354 AAACAGATCAACAGT-----ATATTTCAAGACACAGCTTGGGATACT 1395
 QY 435 LeuTrpProLysGluArgLysArgGlnArgAlaPhePhePheGlyLeuGluLeuIle 454
 DB 1396 TTATGCCACCAGAAAGGAAAGACAGAGAGCATCTTCTTCTTGGTCTTGACACATCA 1455
 QY 455 ValGlnLeuAspIleGluAlaThrArgThrPhePheArgThrPhePheArgLeuProThr 474
 DB 1456 GTTCAATTCGATACCAAGGCAATTTAGAGCTTCTTCGCTACTTCTTCCGCCCTTCCAAA 1515
 QY 475 TrpMetTrpTrpGlyPheLeuGlySerSerLeuSerSerPheAspLeuValLeuPheSer 494
 DB 1516 TGGATGTGGCAAGGGTCTTAGATCAACATTAAACATCAGGAGATCTCTCTTCTTCTTGT 1575
 QY 495 MetTyrMetPheValLeuAlaProAsnSerMetArgMetSerLeuValArgHisLeuLeu 514
 DB 1576 TTATACATGTCGTCATTTCCCAACAAATTTGAGAAAGGCTCTCATCAATCATCTCATC 1635
 QY 515 SerAspProSerGlyAlaValMetValArgAlaTyrLeuGlu 528

[illegible]

QY 397 GluAlaProLysTyrAlaSerValIleAlaLysIleLeuLysGlnAspAsnSerAlaTyr 416
 DB 1290 GAAGCTCCTTAATATGACGAGTAATGCAAGATTTAGGGAAGGAATTCAAAACAG 1349
 QY 417 ValValSerGlyGlnSerSerAlaValAsnIleSerMetGlnAlaTrpSerLeuTrp 436
 DB 1350 ATGCTTGATCATGGAAGATACACCAACATCTCAAGCAAGCTTGGGAACACTTTGG 1409
 QY 437 ProLysGluArgLysArgGlnArgAlaPhePheLeuPheGlyLeuGluLeuIleValGln 456
 DB 1410 CCCTTGAAGGAAGAACAGAGAGATCTTCTCTTTGGATAGACACTGATGTCACG 1469
 QY 457 LeuAspIleGluAlaThrArgThrPhePheArgThrPhePheArgLeuProThrTrpMet 476
 DB 1470 ATGGATATTGAGGGACCCGACATCTTCGGGACATCTTCGCTGCCACATGGATG 1529
 QY 477 TrpTrpGlyPheLeuGlySerSerLeuSerSerPheAspLeuValLeuPheSerMetTyr 496
 DB 1530 TGGTGGGGTCTTCTGGATCTTGGTATCATCACTCACTGACTGATATAATATTGGCTTTAC 1589
 QY 497 MetPheValLeuAlaProAsnSerMetArgMetSerLeuValAlaGlnHisLeuLeuSerAsp 516
 DB 1590 ATGTTTATCATAGCACCAGCATCCCTGAGAAATGGTCTGGTTAGACATTTGCTTCRGAC 1649
 QY 517 ProSerGlyAlaValMetValArgAlaTyrLeu 527
 DB 1650 CCGACAGGAGGAACAATGTAAAGCGTATCTC 1682

RESULT 7:
 AAZ46328
 ID AAZ46328 standard; cdna; 1780 BP.
 AC AAZ46328;
 XX
 DT 07-WAR-2000 (first entry)
 XX
 DE Romaine lettuce epsilon lycopene cyclase cDNA.
 XX
 KW lycopene; epsilon cyclase; bicyclic epsilon carotene; biosynthesis;
 KW carotenoid; epsilon ring; lactucaxanthin; protein production;
 KW transgenic plant; altered synthesis; overexpression;
 KW nutritional value; pharmacology; colour; cancer; antisense;
 KW accumulation; substrate; beta-carotene; vitamin A; ds.
 XX
 OS Lactuca sativa.
 FH
 FX Key Location/Qualifiers
 FT CDS 77..1677
 FT /tag= a
 FT /product= "Romaine lettuce epsilon lycopene cyclase"
 FT /transl_except= (pos:182..187, aa:Arg)
 FT /note= "This codon has an apparent 3 base insertion,
 FT leading to an insertion of an additional Arg
 FT residue in the protein"
 FT unsure 1457..1459
 FT /tag= c
 FT /note= "Encodes Met"
 XX
 XX WO9561399-A1.
 XX
 XX 02-DEC-1999.
 PD
 XX 25-MAY-1999; 99WO-US10461.
 XX
 XX 26-MAY-1998; 98US-0084222.
 XX
 XX (UYMA-) UNIV MARYLAND BALTIMORE.
 PA
 XX Cunnigham FX;
 XX
 XX WPI: 2000-062667/05.
 DR
 DR P-PSDB; AAY52592.

XX
 PT New eukaryotic epsilon-lycopene cyclase and related DNA, used to
 PT regulate carotenoid synthesis in plants and other hosts, e.g. for
 PT treatment of cancer
 XX
 XX Claim 4; Fig 3; 38pp; English.
 XX
 CC This sequence represents cDNA encoding epsilon lycopene cyclase from
 CC romaine lettuce. This enzyme adds two epsilon rings to lycopene to
 CC form bicyclic epsilon-carotene. Carotenoids with two epsilon rings
 CC are found in significant amounts in relatively few plants (romaine
 CC lettuce produces an abundance of such a carotenoid, lactucaxanthin),
 CC while carotenoids with two beta rings are ubiquitous and those
 CC with one beta and one epsilon ring are common. The invention relates to
 CC epsilon lycopene cyclases capable of forming bicyclic epsilon-carotene,
 CC and also to methods of transforming known carotenoids into novel or rare
 CC products. Epsilon lycopene cyclase nucleic acids may be used for
 CC recombinant protein production and for the generation of transgenic
 CC plants that have altered patterns of carotenoid synthesis. Such plants
 CC may accumulate new or rare carotenoids or overexpress known carotenoids,
 CC resulting in altered nutritional value, pharmacology or colour.
 CC Carotenoids isolated from such transgenic plants are useful
 CC therapeutically, e.g., for treatment of cancer. Alternatively, epsilon
 CC lycopene cyclase antisense nucleic acids can be used to cause
 CC accumulation of the substrates of the enzyme, specifically beta-carotene
 CC or derived products such as vitamin A. The isolated eukaryotic enzyme
 CC allows synthesis of epsilon, epsilon-carotene (which can be
 CC isolated from natural sources only in small amounts) on a large scale,
 CC enabling bulk production of derived carotenoids.
 XX
 SQ Sequence 1780 BP; 526 A; 335 C; 382 G; 536 T; 1 other;
 Alignment Scores:
 Pred. No.: 2 18e-195 Length: 1780
 Score: 1931.00 Matches: 375
 Percent Similarity: 78.51% Conservative: 56
 Best Local Similarity: 68.31% Mismatches: 78
 Query Match: 70.40% Indels: 40
 DB: 21 Gaps: 5
 US-09-701-395a-23 (1-529) x AAZ46328 (1-1780)
 QY 1 MetGluLeuLeuGlyValArgAsnLeu-----IleSerSerCysProVal 15
 DB 77 ATGGAGTGTCTTGGAGCTCGAAACATGACGGCAACAATGGCGTTTACGTGCCCTAGA 136
 QY 16 TrpThr-----PheGlyThr 20
 DB 137 TTCACGGACTGTAATATCAGGCACAAATTTTCGTACTGAAACACGAAGATTACTAAT 196
 QY 21 ArgAsnLeuSerSerSerLysLeuAlaTyrAsnIleHisArgTyrGlySerSerCysArg 40
 DB 197 TTATCAGCACTGCTTCTGTCGCTCAATTAAGTCAGCGCTAAAGCGACCGTTGTGTA 256
 QY 41 ValAspPheGlnValArgAlaAspGlySerGlySerArgSerSerValAlaTyrLys 60
 DB 257 GTGGAT-----AAA 265
 QY 61 GluGlyPhe-----ValAspGluGluAspPheIleLysAlaGlyGlySerGluLeuLeu 78
 DB 266 CAAGGGATTTCCGTAGCAGACGAAGAGATTATGTGAAGCGGTGGATCGGAGCTGTTT 325
 QY 79 PheValGlnMetGlnThrLysSerMetGluLysGlnAlaLysLeuAlaAspLysLeu 98
 DB 326 TTTGTTCAATGACGAGGACTAAGTCCATGGAAGCCAGCTTAACCTTCCGAAAGGTA 385
 QY 99 ProProIleProPheGlyGluSerValMetAspLeuValIleGlyCysGlyProAla 118
 DB 386 GCACAGATACCAATTGGAAATTTGCATCTGCTGGTGTGTAATCGGTGTGGCCCTGCT 445
 QY 119 GlyLeuSerLeuAlaAlaGluAlaLysLeuGlyLeuLysValGlyLeuIleGlyPro 138
 DB 446 GGCCTTGTCTTGTGTCAGAGTCAGCCAAACTAGGTTTGAACGTTGGACTCATTTGGCCCT 505

QY 57 -----ValAlaTyrLysGluGlyPheValAspGluAspPheIle 70
Db AATAGCAGACAGTAGTGGCGTAATTCGCCAGAGGATTTTCGGAACGAAGAAGATTTTCATC 437
QY 71 LysAlaGlyCysSerGluLeuPheValGlnMetGlnThrLysSerMetGluLys 90
Db AAGCTGGGTGGTTCGCGAGCTCTTTATGTTCAATGACAGCAATAAAGCTATGGATGTG 497
QY 91 GlnAlaLysLeuAlaAspLysLeuProPheIleProPheGlyGluSerValMetAspLeu 110
Db TACTCCAAAATTTCCGATAGCTGCGTCAAAATATCAGATGCCAATGAACCTGCTGGATATG 557
QY 111 valValIleCysGlyProAlaGlyLeuSerLeuAlaAlaGluAlaLysLeuGly 130
Db GTGGTATTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 617
QY 131 LeuLysValGlyLeuIleGlyProAspLeuProPheThrAsnSerValGlyValTrpGlu 150
Db TTTAAAGTTGGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 677
QY 151 AspGluPheLysAspLeuGlyLeuGluArgCysIleGluHisAlaTrpLysAspThrIle 170
Db GATGAATTCAGAGCATTTGGAGCTGGAGGCTGTATCGAGCACCTTTGGCGTGATACCAT 737
QY 171 ValTyrLeuAspAsnAspAlaProValLeuIleGlyArgAlaTyrGlyArgValSerArg 190
Db GTGTATATTGATGATGACATCTATATATATATATATATATATATATATATATATATAT 797
QY 191 HisLeuLeuHisGluLeuLysArgCysValGluSerGlyValSerTyrLeuAsp 210
Db CAATTAATTCACAGAGCTGGTGGAGCTGGTGGAGCTGGTGGAGCTGGTGGAGCTGGTGG 857
QY 211 SerLysValGluArgIleThrGluAlaGlyAspGlyHisSerLeuValValCysGluAsn 230
Db GCGAAGATGCAAAATATTATGGAAGACCTGATGGACATAGGCTGTGTGCTTGTGAAGCT 917
QY 231 GluIlePheIleProCysArgLeuAlaThrValAlaSerGlyAlaLysLeuLys 250
Db GGTGTCACTATTCCTCCGAGGCTGTAACCTGTTGCACTGTGGAGCAGCTTCAGGGAACT 977
QY 251 LeuGluTyrGluValGlyProArgValCysValGlnThrAlaTyrGlyValGluVal 270
Db CTGGAGTATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1037
QY 271 GluValGluAsnAsnProTyrAspProAsnLeuMetValPheMetAspTyrArgAspTyr 290
Db GAGGTGGAACACAGCTTATGATCCCAATGTGATGTGTGATGGACTACAGAGACTAC 1097
QY 291 MetGlnGlnLysLeuGlnCysSerGluGluGluTyrProThrPheLeuTyrValMetPro 310
Db ACTAAACTGAGCGTTCAATCTCTGGAGGCAAGTATCCAACTTCTGTATGCAATGCCG 1157
QY 311 MetSerProThrArgLeuPheGluPheGluThrCysLeuAlaSerLysAspAlaMetPro 330
Db ATATACCAACTAGGATCTCTTTGAGGAGACTTGTGCTTGCATGATGATGCAATGCC 1217
QY 331 PheAspLeuLysArgLysLeuMetSerArgLeuLysThrLeuGlyIleGlnValThr 350
Db TTTGACCTGCTCAAGAAAAAGCTTATGACAAGATTAACAACATATGCTGTGCTATCAC 1277
QY 351 LysValTyrGluGluGluTrpSerTyrIleProValGlySerLeuProAsnThrGlu 370
Db AAAATATATGAGAGAGGTGGTCTTATATACCTGTTGGTGGTGGTGGTGGTGGTGGTGG 1337
QY 371 GlnLysAsnLeuAlaPheGlyAlaAlaAlaSerMetValHisProAlaThrGlyTyrSer 390
Db CAAGAAGAACTTGCATTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1397
QY 391 ValValArgSerLeuSerGluAlaProLysTyrAlaSerValIleAlaLysIleLeuLys 410
Db GTCGTGAGTACACTGTGACAGAGCTCCAAAGATGCTCTGCAATTCGCAACTGATCAAG 1457
QY 411 GlnAspAsnSerAlaTyrValValSerGlyGlnSerSerAlaValAsnIleSerMetGln 430

Db 1458 AATGACCTGTCAAAAAATGCAATATTCCTCAGAGGAGTGTGGGAATATCTCAATGCAA 1517
QY 431 AlaTrpSerSerLeuTrpProLysGluArgLysArgGlnArgAlaPhePheLeuPheGly 450
Db GCCTGGGAATACTCTTGGCCACAAGAAAGAACGTCAGAGAGCATTTCTCTGTTCGGA 1577
QY 451 LeuGluLeuIleValGlnLeuAspIleGluAlaThrArgThrPheArgThrPhePhe 470
Db CTATCACTTATAGTCCAGCTTGATATTGAGGTATCAGGACATTTCTCCGACCTTCTC 1637
QY 471 ArgLeuProThrTrpMetTrpTrpGlyPheLeuGlySerSerLeuSerSerPheAspLeu 490
Db CGAGTGCACAAATGATGTGGGAGGATTCCTCGGTTCTTAATCTCTCTACGTGATCTC 1697
QY 491 ValLeuPheSerMetTyrMetPheValLeuAlaProAsnSerMetArgMetSerLeuVal 510
Db ATATGTTGGCTTTTATATATCTTCTTATGCTCCGAATGACTTGAGATGGGTCTTATA 1757
QY 511 ArgHisLeuLeuSerAspProSerGlyAlaValMetValArgAlaTyrLeu 527
Db AGGCATCTACTATCTGATCTACAGGGGCGACCATGATAAGAACGTACATA 1808
RESULT 9
AAZ45586
ID AAZ45586 standard; cDNA; 1377 BP.
XX
AC AAZ45586;
XX
DT 06-APR-2000 (first entry)
XX
DE cDNA encoding a potato lycopene epsilon-cyclase protein.
XX
KW Lycopene epsilon-cyclase; carotenoid biosynthesis pathway;
KW isopentenyl pyrophosphate isomerase; beta-carotene hydroxylase;
KW dimethylallyl pyrophosphate; transgenic plant; carotenoid content;
XX lethal photooxidative damage; colouring; ss.
OS Solanum tuberosum.
XX
FH Key Location/Qualifiers
FT CDS 14..1150
FT /*cag= a /product= "lycopene epsilon-cyclase"
FT WO9963055-A1.
XX
PD 09-DEC-1999.
XX
PF 02-JUN-1999; 99WO-US12121.
XX
PR 02-JUN-1998; 98US-0088724.
PR 02-JUN-1998; 98US-0088725.
XX
PA (DYMA-) UNIV MARYLAND BALTIMORE.
XX
PI Cunnigham Fx, Sun Z;
XX
XX WPI; 2000-116366/10.
DR P-PSDB; AAY54291.
XX
XX New carotenoid biosynthesis genes used to provide transgenic plants
PT with altered carotenoid content
XX
XX Disclosure; Fig 15A; 134pp; English.
XX
CC The present sequence encodes a lycopene epsilon-cyclase protein.
CC The enzyme is involved in the carotenoid biosynthesis pathway.
CC specification also describes other enzymes involved in this pathway,
CC such as epsilon-cyclase, isopentenyl pyrophosphate (IPP) isomerase and
CC beta-carotene hydroxylase. IPP isomerase catalyses the reversible
CC conversion of IPP to dimethylallyl pyrophosphate (DMAPP). Beta-carotene
CC hydroxylase is responsible for hydroxylating the beta-endgroup in

CC carotenoids. Lycopene epsilon-cyclase is responsible for the formation
 CC of epsilon-endgroups in carotenoids. The polynucleotides of the
 CC invention are used to provide transgenic plants with altered
 CC carotenoid content. These plants may, for example, may be protected
 CC from lethal photooxidative damage or have altered colouring.

XX
 SQ Sequence 1377 BP; 386 A; 255 C; 297 G; 431 T; 8 other;

Alignment Scores:
 Pred. No.: 1.32e-151 Length: 1377
 Score: 1518.50 Matches: 281
 Percent Similarity: 87.27% Conservativity: 48
 Best Local Similarity: 74.54% Mismatches: 47
 Query Match: 55.36% Indels: 1
 DB: 21 Gaps: 1

US-09-701-395A-23 (1-529) x AA245586 (1-1377)

Qy 151 AspGluPheLysAspLeuGluArgCysLeuGluHisAlaTrpLysAspThrIle 170
 Db 14 GATGAGTTCAAGATCTTGGTCTTCAAGCCGTGATGAACATGTTTGGCGGATACCAATT 73
 Qy 171 ValTrpLeuAspAsnAlaProValLeuIleGlyArgAlaTrpGlyArgValSerArg 190
 Db 74 GTATATCTTGATGATGATGATCTTATCTTATGCGCGCTATGGAAGAGTTAGTCGC 133
 Qy 191 HisLeuLeuHisGluGluLeuLysArgCysValGluSerGlyValSerTrpLeuAsp 210
 Db 134 CATTTACTGCACGAGGAGTTACTCAAAAGGTGTGGAGGAGGAGTTTGTATCTAATAC 193
 Qy 211 SerLysValGluArgIleThrGluAlaGlyAspGlyHisSerLeuValValCysGluAsn 230
 Db 194 TCGAAGTGGATAGATTGTTGAGGCCCAAAATGCCACAGCTTGTAGAGTGGCAGGGT 253
 Qy 231 GluIlePheIleProCysArgLeuAlaThrValAlaSerGlyAlaAlaSerGlyLysLeu 250
 Db 254 GATGTTGTGATCCCTCGAGGTTGTGACTGTTCGCGGAGCAGCTCGGGGAAATTC 313
 Qy 251 LeuGluTrpGluValGlyGlyProArgValCysValGluThrAlaTrpGlyValGluVal 270
 Db 314 TTGCAGTATGATGTTGGGAGGTCCTAGAGTTCTGTTCAACAGCTTATGGAGTGAAGTT 373
 Qy 271 GluValGluAsnAsnProTrpAspProAsnLeuMetValPheMetAspTrpArgAspTrp 290
 Db 374 GAGTCTGATCAACATCCATTGTCGCGGAGCCTGATGTTTTCATGGATTATAGACTAT 433
 Qy 291 MetGluGluLysLeuGluCysSerGluGluTrpProThrPheLeuTrpValMetPro 310
 Db 434 GTCAGACACGCGCTCAATCTTTAGAGCTAAATATCCACATTTCTATGCCATGCC 493
 Qy 311 MetSerProThrArgLeuPheGluGluThrCysLeuAlaSerLysAspAlaMetPro 330
 Db 494 ATGCTCCACACGAGTCTTTTCGAGGAAACTGTTGGCTTCAAAAGATGCAATGCCA 553
 Qy 331 PheAspLeuLeuLysArgLysLeuMetSerArgLeuLysThrLeuGlyIleGluValThr 350
 Db 554 TTCGATCTGTTAAAGAAATAATGATGTTAGATTTGACACCCCTCGGTGTAAGAATAAA 613
 Qy 351 LysValTrpGluGluTrpSerTrpIleProValGlyGlySerLeuProAsnThrGlu 370
 Db 614 GAAATTTATGAGGAGGATGCTTACATACACAGTTGGAGGATCTTGGCCAAATACAGAA 673
 Qy 371 GlnLysAsnLeuAlaPheGlyAlaAlaAlaSerMetValHisProAlaThrGlyTrpSer 390
 Db 674 CAAAAACACCTTGCATTGTTGGTGTCTGCTAGCTAGCTATCCAGCCACAGGTTATTCA 733
 Qy 391 ValValArgSerLeuSerGluAlaProLysTrpAlaSerValIleAlaLysIleLeuLys 410
 Db 734 GTCGTGATGATCCTGCTGAGGCTCCAAAATGCGCCTCGCTGCAAAATATATATACGA 793
 Qy 411 GlnAspAsnSerAlaTrpValValSerGlyGlnSerSerAlaValAlaSerMetGln 430
 Db 794 CAAATCATACGAATATGCTTACT--AGTTCAAGTACCCCGAGTATTTCAACTCAA 850

Qy 431 AlaTrpSerLeuTrpProLysGluArgLysArgGlnArgAlaPhePheLeuPheGly 450
 Db 851 GCTTGGACACTCTTGGCCACAGACGAAACACAGAAAGTCTGTTTCTTATTTGA 910
 Qy 451 LeuGluLeuIleValGlnLeuAspIleGluAlaThrArgThrPheArgThrPhePhe 470
 Db 911 CTGGCTCTGATATTGACATGATATTGAGGGGATAAGTCAATTTTCCGGCGGTTCTTC 970
 Qy 471 ArgLeuProThrTrpMetTrpTrpGlyPheLeuGlySerSerLeuSerSerPheAspLeu 490
 Db 971 CGTGTGCCAAAATGATGTGCGAGGATTTCTTGGTTCAAGTCTTCTTNAGCAGACCTC 1030
 Qy 491 ValLeuPheSerMetTrpMetPheValLeuAlaProAsnSerMetArgMetSerLeuVal 510
 Db 1031 ATGTTATTTGCTTACATGTTTATTTATTCACCAATGACATGAGAGGCTTATTC 1090
 Qy 511 ArgHisLeuLeuSerAspProSerGlyAlaValMetValArgAlaTrpLeu 527
 Db 1091 AGACATCTTTTATCTGATCTTACTGTGCTCAACATTTGATAAGAACTTATCTT 1141

RESULT 10

AAV19124 standard; cDNA; 1272 BP.

XX AAV19124;

XX AC 10-AUG-1998 (first entry)

XX DE Nucleotide sequence of epsilon cyclase cDNA clone 7-6.

XX XX Epsilon cyclase; SSU/crtB fusion protein; carotenoid; lycopene; lutein;
 KW zeaxanthin; canthaxanthin; vitamin; colourant; ss.

XX OS Brassica napus.

XX PN W09806862-A1.

XX PD 19-FEB-1998.

XX PF 08-AUG-1997; 97WO-US14035.

XX PR 09-AUG-1996; 96US-0024145.

XX PA (CALJ) CALGENE INC.

XX PI Shewmaker CK;

XX DR WPI; 1998-159551/14.

XX PT Use of constructs comprising a carotenoid biosynthesis gene - for
 PT producing plants and seeds having altered carotenoid levels,
 PT modified fatty acid compositions or altered tocopherol levels.

XX PS Example 1B; Fig 10; 70pp; English.

XX CC This is the nucleotide sequence of the Brassica napus epsilon cyclase
 CC cDNA clone 7-6. It was used in the method of the invention with
 CC the SSU/crtB fusion protein (AAV19122), to create plants and seeds with
 CC altered carotenoid levels. They can be used for the production of
 CC particular carotenoids, e.g. alpha-carotene, beta-carotene, lycopene,
 CC lutein, zeaxanthin, canthaxanthin, which can be used as supplements,
 CC particularly vitamin supplements, as vegetable oil based food products,
 CC and food ingredients, as feed additives in animal feeds, as colourants
 CC or therapeutic agents.

XX SQ Sequence 1272 BP; 342 A; 271 C; 311 G; 348 T; 0 other;

Alignment Scores:

Pred. No.: 2.15e-138 Length: 1272
 Score: 1394.00 Matches: 265
 Percent Similarity: 85.51% Conservativity: 30
 Best Local Similarity: 76.81% Mismatches: 46

Query Match:	50.82%	Indels:	4
DB:	19	Gaps:	2
US-09-701-395A-23 (1-529) x AAV19124 (1-1272)			
Qy	106	SerValMetAspLeu-----ValValIleGlyCysGlyProAlaGlyLeuSerLeuAla	123
Db	67	AGTGTGATGGATATCTGCAGAAATCGGCTGTGTTGTGCTCTGCTGGTTTACGCTTGCGC	126
Qy	124	AlaGluAlaAlaLysLeuGlyLeuLysValGlyLeuIleGlyProAspLeuProPheThr	143
Db	127	GCTGAATCAGCTAAGTTAGACTTAAAGTTGGACTGATTGGTCTCTGACATCTTCCCTTCACT	186
Qy	144	AsnAsnTyrGlyValTrpGluAspGluPheLysAspLeuGlyLeuGluArgCysIleGlu	163
Db	187	ARCAACTACGGTGTGTTGGGAAGATGAGTTCACAGCATCTGGCTTGGCTTGCACAAATGATTTCAG	246
Qy	164	HisAlaTrpLysAspThrIleValTyrLeuAspAsnAlaProValLeuIleGlyArg	183
Db	247	CATGTTTGGAGAGATACCTTGTGTATCTGGACGATGACAAATCTTATACCATTTGGTTCGT	306
Qy	184	AlaTyrGlyArgValSerArgHisLeuLeuHisGluGluLeuLeuLysArgCysValGlu	203
Db	307	GCTTATGGAAGAGTTAGTCCAGCTTTACTTTCACGAGGAGTTCTTCTGAGGAGGTGTGGAG	366
Qy	204	SerGlyValSerTyrLeuAspSerLysValGluArgIleThrGluAlaGlyAspGlyHis	223
Db	367	TCAGGTGTCTCGTATCTTACGTCCTCAAGTTGAGACATCAACAGAAAGCTCTGATGGCGCTT	426
Qy	224	SerLeuValValCysGluAsnGluIlePheIleProCysArgLeuAlaThrValAlaSer	243
Db	427	AGCCTGTTTCTGTGTGAACAAACACCCCTGTGTTCTCGACGGCTTCCACGCTGCTCTCT	486
Qy	244	GlyAlaAlaSerGlyLysLeuLeuGluTyrGluValGlyGlyProArgValCysValGln	263
Db	487	GGAGCAGCTTCTGGGAAGCTCTTGCAATACGAAGTTGAGGGCCCTAGAGTCTGTGTCCAA	546
Qy	264	ThrAlaTyrGlyValGluValGluAsnAsnProTyrAspProAsnLeuMetVal	283
Db	547	ACTGCTTACGGCTTGGAGGTGAGGTGGAAGAGATCCATATGATPCAGAGACGATGGTG	606
Qy	284	PheMetAspTyrArgAspTyrMetGlnGlnLysLeuGlnCysSerGluGluGluTyrPro	303
Db	607	TTCATGGATTACAGAGATTATACAAACGAGAAATCCGAGCTTAGAAGCTGAATATCCA	666
Qy	304	ThrPheLeuTyrValMetProMetSerProThrArgLeuPheGluGluThrCysLeu	323
Db	667	ACGTTTCTCTACGCCATGCCTATGACAAAGACCCAGAGTCTTCTTTGAGGAGACATGTCTT	726
Qy	324	AlaSerLysAspAlaMetProPheAspLeuLeuLysArgLysLeuMetSerArgLeuLys	343
Db	727	GCTTCAAAAGATGTCATGCCCTTGTATTTGCTTAAAAAGAAAGCTCTTCTTTCAGATTAAG	786
Qy	344	ThrLeuGlyIleGlnValThrLysValTyrGluGluGluTrpSerTyrIleProValGly	363
Db	787	ACACTCGGAATCCGAATATAAGACTTACGAAGAGGAATGGTCTTATATCCAGTAGGT	846
Qy	364	GlySerLeuProAsnThrGluGlnLysAsnLeuAlaPheGlyAlaAlaSerMetVal	383
Db	847	GGTTCCTTGCACAAACGGAACAAAAGAAATCTCGGCTTTGGTGTGTCAGCTAGCATGGTT	906
Qy	384	HisProAlaThrGlyTyrSerValValArgSerLeuSerGluAlaProLysTyrAlaSer	403
Db	907	CATCCTTCACACAGGCTATTTCAGTTGTGAGATCTTGTCTGAAGCTTCCAAAATACGCATCA	966
Qy	404	ValIleAlaLysIleLeuLysGlnAspAsnSerAlaTyrValValSerGlyGlnSerSer	423
Db	967	GTATCGCTAAATATATAAACAATGAGACCACTACTCTCTTCACC-----AGACACATC	1020
Qy	424	AlaValAsnIleSerMetGlnAlaTrpSerSerLeuTrpProLysGluArgLysArgGln	443
Db	1021	AACACCAATATTCAAGACAAGCTTGGATATCTTTATGGCCACACAGAAAGAAACGACAG	1080

QY	106	SerValMetAspLeu-----ValValIleGlyCysGlyProAlaGlyLeuSerLeuAla	123
DB	67	AGTGTGATGGATATCTCGAAGTTCGGCTTGTGTGGTCTGTGGTTAGCCTTGCG	136
QY	124	AlaGluAlaAlaLeuLeuGlyLeuLysValGlyLeuIleGlyProAspLeuProPheThr	143
DB	127	GCTGAATCAGCTAAAGTTAGGACTTAAAGTTGGACTGATGTGCTCTGACCTTCCTTCACT	186
QY	144	AsnAsnTyrGlyValTrrpGluAspGluPheLysAspLeuGlyLeuGluAtrGcysIleGlu	163
DB	187	AACAACACAGCGTGTGGGAGATGAGTTCAACAGATCTTGGCTTGGCTTGGCAAAATGATTGAG	246
QY	164	HisAlaTrpLysAspThrIleValTyrLeuAspAsnAlaProValLeuIleGlyArg	183
DB	247	CATGTTTGGAGATACCTCTGTATCTGGACGATGACATCTATTACCATTTGGTCTGT	306
QY	184	AlaTyrGlyArgValSerArgHisLeuLeuHisGluGluLeuLeuLysArgCysValGlu	203
DB	307	GCTTATGGAGACATTAGTCGACGTTTACTTCCAGGAGTCTTGTAGGAGGTGTGTGGAG	366
QY	204	SerGlyValSerTyrLeuAspSerLysValGluArgIleThrGluAlaGlyAspGlyHis	223
DB	367	TCAGGTGTCTCGTATCTTAGCTCCAAAGTTCGAGACATACAGAAGCTCCTCATGGCCCT	426
QY	224	SerLeuValValCysGluAsnGluIlePheIleProCysArgLeuAlaThrValAlaSer	243
DB	427	AGCCTTGTTCCTGTGTAACAAACACCCCTTGTCCGTGACGGCTTGCCTGTCTTCT	486
QY	244	GlyAlaAlaSerGlyLysLeuLeuGluTyrGluValGlyGlyProArgValCysValGln	263
DB	487	GGAGCAGCTCTGGGAAGCTCTCCATACGAAGTTGGAGGGCCTAGAGTCTGTGTCCAA	546
QY	264	ThrAlaTyrGlyValGluValGluValGluAsnProTyrAspProAsnLeuMetVal	283
DB	547	ACTGCTTACGGCTTGGAGTGTAGGTGGAAAGAGTCCATATGATCCAGACCATGGTG	606
QY	284	PheMetAspTyrArgAspTyrMetGlnLysLeuGlnCysSerGluGluGluTyrPro	303
DB	607	TTCATGTGATTACAGAGATTATACAAACGAGAATAATCCCGAGCTTAGAGCTGAATATCCA	666
QY	304	ThrPheLeuTyrValMetProMetSerProThrArgLeuPheGluGluThrCysLeu	323
DB	667	ACGTTTCTCTACGCCATTCGCTATGACAAAGACACGAGTCTTCTTTGAGGACATGTCTT	726
QY	324	AlaSerLysAspAlaMetProPheAspLeuLeuLysArgLysLeuMetSerArgLeuLys	343
DB	727	GCTTCAAAAGATGTCATGCCCTTGTATTGCTTAAAAAAGACGCTCTGTGTGAGATTAGAG	786
QY	344	ThrLeuGlyIleGlnValThrLysValTyrGluGluGluTrrpSerTyrIleProValGly	363
DB	787	ACACTCGGAATCCGAATACTAAAGACTTACGAAGAGGAATGGTCTTATATCCCATGTGT	846
QY	364	GlySerLeuProAsnThrGluGlnLysAsnLeuAlaPheGlyAlaAlaSerMetVal	383
DB	847	GGTTCCTTGCCAAACAGGAACAAAGAATCTCGCTTTGGTGTGTCAGCTAGCATGGTT	906
QY	384	HisProAlaThrGlyTyrSerValValArgSerLeuSerGluAlaProLysTyrAlaSer	403
DB	907	CATCTCGCAACAGGCTATTCACTGTGAGATCTTGTCTGAAGCTCCAAATATACCATCA	966
QY	404	ValIleAlaLysIleLeuLysGlnAspAsnSerAlaTyrValValSerGlyGlnSerSer	423
DB	967	GTCTATCGCTAAATATACTAAACATGAGACCATCTACTTCTCTCAC-----AGACACATC	1020
QY	424	AlaValAsnIleSerMetGlnAlaTrrpSerSerLeuTrrpProLysGluArgGln	443
DB	1021	AACACCAATATTCAAGACAAGCTTGGATACTTTTATGCCACCAGAAAGGAACGACAG	1080
QY	444	ArgAlaPheLeu	448
DB	1081	AGAGCATTTCTTCTA	1095

```
Db 306 GCTTATGGAAGAGTTAGTCCAGCTTTACTTCACGAGGAGCTCTGAGGAGGTGTGTGGAG 365
Qy 204 SerGlyValSerTyrLeuAspSerLysValGluArgIleThrGluAlaGlyAspGlyHis 223
Db 366 TCAGGTGTCCTCGTATCTTACCTCCAAAGTTGAGAGATACAGAGCTCTGATGGCCTT 425
Qy 224 SerLeuValValCysGluAsnGluIlePheIleProCysArgLeuAlaThrValAlaSer 243
Db 426 AGGCTTGTTCCTCGTCAACAAACACCCCTGTTCCGTCGACAGGCTTGCCACTGTTGCTTCT 485
Qy 244 GlyAlaAlaSerGlyLysLeuLeuGluTyrGluValGlyClyProArgValCysValGln 263
Db 486 GGAGCAGCTTCTGGGAAGCTCTTGCATACGAAGTTGGAGGGCTAGAGTCTGTGTCCAA 545
Qy 264 ThrAlaTyrGlyValGluValGluValGluAsnAsnProTyrAspProAsnLeuMetVal 283
Db 546 ACTGCTTACGCTTGGAGGTGAGGTGGAGTGGAAAGAGTCCATATGATCCAGAGCAGATGGTG 605
Qy 284 PheMetAspTyrArgAspTyrMetGlnGlnLysLeuGlnCysSerGluGluGluTyrPro 303
Db 606 TTCATGGATTACAGAGATTATACAAACGAGAGAAATCCGGAGCTTAGAAGCTGAATATCCA 665
Qy 304 ThrPheLeuTyrValMetProMetSerProThrArgLeuPheGluGluThrCysLeu 323
Db 666 AGCTTCTCTACGCCCTGCTATGACAAAGACACGAGAGCTCTTCTTGGAGGAGATGTCTT 725
Qy 324 AlaSerLysAspAlaMetProPheAspLeuLeuLysArgLysLeuMetSerArgLeuLys 343
Db 726 GCTTCAAAAGATCTATGCGCTTGTATGCTTAAAGAAAGCTCTGTGTTGAGATTAGAG 785
Qy 344 ThrLeuGlyIleGlnValThrLysValTyrGluGluGluTyrSerTyrIleProValGly 363
Db 786 ACACCTCGGAATCGAATACTAAAGACTTACGAAGAGAGTGTCTTATATATCCAGTAGGT 845
Qy 364 GlySerLeuProAsnThrGluGlnLysAsnLeuAlaPheGlyAlaAlaSerMetVal 383
Db 846 GGTTCCTTCCCAACACGGAACAAAGAAATCTCGCCTTTGGCGCTGCAGCTAGCATGGTA 905
Qy 384 HisProAlaThr 387
Db 906 CATCCGCGCAACA 917

RESULT 13
AAX25064
ID AAX25064 standard; DNA; 962 BP.
AC AAX25064;
XX
DT 05-JUL-1999 (first entry)
XX
DE Brassica napus lycopene epsilon cyclase cDNA clone 9-4.
XX
KW Lycopene epsilon cyclase; phytoene synthase; crtB gene; carotenoid;
KW xanthophyll; pigment; beta-carotene; alpha-carotene; phytoene;
KW transgenic plant; oilseed plant; seed oil; antisense; ss.
XX
OS Brassica napus.
XX
PN WO9907867-A1.
XX
PD 18-FEB-1999.
XX
PF 06-AUG-1998; 98WO-US16466.
XX
PR 08-AUG-1997; 97US-0908758.
XX
PA (CALJ ) CALGENE LLC.
XX
PI Shewmaker CK;
XX
DR WPI; 1999-180495/15.
XX
```

```
PT Altering xanthophyll content of seeds by transformation - used to
PT produce seed oils of increased carotenoid content, e.g. Brassica and
PT cotton
XX
PS Example 1; Fig 9; 92pp; English.
XX
CC This is the nucleotide sequence of Brassica napus lycopene
CC epsilon cyclase cDNA clone 9-4, which was obtained by PCR using
CC primers designed from an Arabidopsis epsilon cyclase gene. The
CC xanthophyll content of seeds is altered by transforming cells of a
CC host plant with at least one construct containing: (i) a transcription
CC initiation region of a gene preferentially expressed in seeds; (ii)
CC the sequence for a plastid transit peptide; (iii) DNA from a
CC carotenoid synthesis gene coding region, and (iv) a transcription
CC terminator. The transformed cells are regenerated to plants and
CC these, or their progeny, grown to produce seeds. The method is
CC especially used to increase the carotenoid content in oilseed
CC plants. Expression of phytoene synthase, encoded by the crtB gene
CC (see AAX25063), together with an antisense epsilon cyclase gene
CC results in large increases in levels of alpha-carotene,
CC beta-carotene and phytoene, and an alteration of alpha- to
CC beta-carotene ratio, in seeds of transformed plants.
XX
```

SQ Sequence 962 BP; 247 A; 203 C; 248 G; 264 T; 0 other;

Alignment Scores:

Pred. No.:	4.04e-118	Length:	962
Score:	1202.00	Matches:	224
Percent Similarity:	88.38%	Conservative:	27
Best Local Similarity:	78.87%	Mismatches:	31
Query Match:	43.82%	Indels:	2
DB:	20	Gaps:	1

US-09-701-395a-23 (1-529) x AAX25064 (1-962)

```
Qy 106 SerValMetAspLeu-----valValIleGlyCysGlyProAlaGlyLeuSerLeuAla 123
Db 66 AGTGTGATGGATATCTGCAGAAATTCGGCTTGTGTGCTCTCTGCTGGTTAGCCTTGGOT 125
Qy 124 AlaGluAlaAlaLysLeuGlyLysValGlyLeuIleGlyProAspLeuProPheThr 143
Db 126 GCAGAAATCAGCAAGGTAGGTCTCAAGTTGGATCATTTGGTCTGTGATCTCTCTTCACT 185
Qy 144 AsnAsnTyrGlyValTyrGluAspGluPheLysAspLeuGlyLeuGluArgCysIleGlu 163
Db 186 AACAACTACGGTGTGGGAAGATGAGTTCAACGATCTTGGCTTGCAAAATGATTATTCAG 245
Qy 164 HisAlaTrpLysAspThrIleValTyrLeuAspAsnAspAlaProValLeuIleGlyArg 183
Db 246 CATGTTTGGAGAGATACCTTGTGTATCTGTGACGATGACAATCTATTACCATTTGGTCTGT 305
Qy 184 AlaTyrGlyArgValSerArgHisLeuLeuHisGluGluLeuLeuLysArgCysValGlu 203
Db 306 GCTTATGGAGAGATTAGTCGAGCTTACTTCACGAGGAGCTCTTGAGAGAGTGTGTGGAG 365
Qy 204 SerGlyValSerTyrLeuAspSerLysValGluArgIleThrGluAlaGlyAspGlyHis 223
Db 366 TCAGGTGTCGTATCTTAGCTCCAAAGTTTCAGAGACGACAGAGCTCTGTATGGCCTT 425
Qy 224 SerLeuValValCysGluAsnGluIlePheIleProCysArgLeuAlaThrValAlaSer 243
Db 426 AGGCTTGTTCCTGTGAACAAACACCCCTTGTTCGTCGACGCTTGGCACCTGTCTCTCT 485
Qy 244 GlyAlaAlaSerGlyLysLeuLeuGluTyrGluValGlyGlyProArgValCysValGln 263
Db 486 GGAGCAGCTTCTGGGAAGCTCTTGCATACGAAGTTGGAGGGCTTAGAGTCTGTGTCCAA 545
Qy 264 ThrAlaTyrGlyValGluValGluValGluAsnAsnProTyrAspProAsnLeuMetVal 283
Db 546 ACTGCTTACGCTTGGAGGTGAGGTGGAGAAAGAGTCCATATGATCCAGAGCAGATGGTG 605
Qy 284 PheMetAspTyrArgAspTyrMetGlnGlnLysLeuGlnCysSerGluGluGluTyrPro 303
Db 906 CATCCGCGCAACA 917
```

Db 606 TTCATGGATTACAGATTATACAAACGAGAAATCCGGAGCTTAGAAGCTGAATATCCA 665
 Qy 304 ThrPheLeuTyrrValMetProMetSerProThrArgLeuPheGluGluThrCysLeu 323
 Db 666 AGCTTCTCTACGCCATGCTATGACAAAGACAGAGCTTCTTTGAGGAGACATGCTT 725
 Qy 324 AlaSerLysAspAlaMetProPheAspLeuLeuLysArgLysLeuMetSerArgLeuLys 343
 Db 726 GCTTCAAAAGATGTCATGCCCTTTGATTGCTTAAAGAGAGCTCTTGTGAGATTAG 785
 Qy 344 ThrLeuGlyIleGlnValThrLysValTyrrGluGluGluTyrSerTyrrIleProValGly 363
 Db 786 ACACCTCGAATCCGATACTAAAGACTTACGAAGAGGATGCTCTATATCCAGTAGGT 845
 Qy 364 GlySerLeuProAsnThrGluGlnLysAsnLeuAlaPheGlyAlaAlaSerMetVal 383
 Db 846 GGTTCCTTCCCAACACGGAACAAAGAAATCTCGCCTTTGGCGTGCAGCTAGCATGTA 905
 Qy 384 HisProAlaThr 387
 Db 906 CATCCCGCAACA 917

RESULT 14

ABA97355

ID ABA97355 standard; cDNA; 962 BP.

AC ABA97355;

DT 12-APR-2002 (first entry)

DE Brassica napus epsilon cyclase cDNA clone 9-4.

KW Carotenoid compound; oilseed; plant; vitamin A; ss; gene.

OS Brassica napus.

PN WO200188169-A2.

PD 22-NOV-2001.

PF 11-MAY-2001; 2001WO-US15264.

PR 12-MAY-2000; 2000US-0570140.

PX (MONS) MONSANTO TECHNOLOGY LLC.

PY Shewmaker CK;

PI WPI; 2002-075318/10.

XX Altering carotenoid level/content in oilseed plants, by transforming
 PT plants with construct having DNA sequence encoding a carotenoid
 PT biosynthesis gene, plastid transit peptide, transcription initiator and
 PT terminator

PS Disclosure; Page 71; 114pp; English.

XX This invention relates to altering the carotenoid content or
 CC composition in seed from a host plant. This is achieved by transforming
 CC cells of host plant with a construct of operably linked components, a
 CC transcriptional initiation region from a gene expressed in a plant seed,
 CC a plastid transit peptide, a nucleic acid sequence encoding a carotenoid
 CC biosynthesis gene from eukaryotic source, and a transcriptional
 CC termination region. The method is useful for increasing carotenoid
 CC content in seed and endosperm of a host plant and altering
 CC carotenoid composition in a plant seed. The method is also useful for
 CC screening transformed corn seeds or transformed endosperms, where the
 CC transformed seed and endosperms are visually determined and selected
 CC based on yellow, orange or red colour as result of the increased
 CC carotenoid content. The transformed seeds provide a source of
 CC modified oils and the oil extracted from the seeds is useful as a food
 CC colourant, or as a food oil with high alpha and beta-carotene levels for
 CC prevention of vitamin A deficiency which can result in night blindness.

CC The level of lutein is also increased in seeds of transformed plants.
 CC The modified oil obtained from the transformed seeds are more oxidatively
 CC stable than naturally occurring oils and are low saturate, high oleic and
 CC low linolenic. This sequence represents the nucleotide sequence of
 CC the Brassica napus epsilon cyclase cDNA clone 9-4.

XX SQ Sequence 962 BP; 247 A; 203 C; 248 G; 264 T; 0 other;

Alignment Scores:

Pred. No.:	4,04e-118	Length:	962
Score:	1202.00	Matches:	224
Percent Similarity:	88.38%	Conservative:	27
Best Local Similarity:	78.87%	Mismatches:	31
Query Match:	43.82%	Indels:	2
DB:	24	Gaps:	1

US-09-701-395A-23 (1-529) x ABA97355 (1-962)

Qy 106 SerValMetAspLeu-----ValValIleGlyCysGlyProAlaGlyLeuSerLeuAla 123
 Db 66 AGTGTGATGGATATCTGCAGAAATTCGGCTTGTGTGGTCTGCTGGTTAGCCTTGGCT 125
 Qy 124 AlaGluAlaAlaLysLeuGlyLeuLysValGlyLeuIleGlyProAspLeuProPheThr 143
 Db 126 GCAGAAATCAGCAAGTTAGGTCTCAAGTTGGACTCATTTGGTCTGATCTTCTTCTACT 185
 Qy 144 AsnAsnTyrrGlyValTrpLysAspGluPheLysAspLeuGlyLeuGluArgCysIleGlu 163
 Db 186 AACAACTACGGTGTGGGAAGATGAGTTCAACGATCTTGGCTTGCAGAAATGATTAG 245
 Qy 164 HisAlaTrpLysAspThrIleValTyrrLeuAspAsnAlaProValLeuIleGlyArg 183
 Db 246 CATGTTTGGAGAGATACCTTTGTATCTGGACGATGACAATCCTATTACCATTTGGCT 305
 Qy 184 AlaTyrrGlyArgValSerArgHisLeuLeuHisGluGluLeuLeuLysArgCysValGlu 203
 Db 306 GCTTATGGAGAGTAGTAGTACGAGCTTACTTACGAGGAGCTTCTTGGAGAGGTGTGGAG 365
 Qy 204 SerGlyValSerTyrrLeuAspSerLysValGluArgIleThrGluAlaGlyAspGlyHis 223
 Db 366 TCAGTGTCTCGTATCTTAGCTCCAAAGTTGAGAGCATAACAGAGCTTCTTGGTCTTCT 425
 Qy 224 SerLeuValValCysGluAsnGluIlePheIleProCysArgLeuAlaThrValAlaSer 243
 Db 426 AGGCTTGTCTTCTGTGAACAAACACCCCTTGTCTCGTACAGCTTGCACCTTCTTCT 485
 Qy 244 GlyAlaAlaSerGlyLysLeuGluTyrrGlyValGlyGlyProArgValCysValGln 263
 Db 486 GGACAGCTTCTGGAGAGCTCTTGCAATACGAAGTTGGAGGCTTAGAGTCTGTGTCCAA 545
 Qy 264 ThrAlaTyrrGlyValGluValGluValGluAsnAsnProTyrrAspProAsnLeuMetVal 283
 Db 546 ACTGCTTACGGCTTGGAGTTGAGTGGAAAGAGTCCATATGATCCAGAGCAGATGGTG 605
 Qy 284 PheMetAspTyrrArgAspTyrrMetGlnLysLeuGlnCysSerGluGluGluTyrrPro 303
 Db 606 TTCATGGATTACAGAGATTATACAAACGAGAAATCCGGAGCTTAGAAGCTGAATATCCA 665
 Qy 304 ThrPheLeuTyrrValMetProMetSerProThrArgLeuPheGluGluThrCysLeu 323
 Db 666 ACGTTTCTTACGCCATGCTTATGACAAAGACACAGAGCTTCTTGTGAGGAGACATGCTT 725
 Qy 324 AlaSerLysAspAlaMetProPheAspLeuLeuLysArgLysLeuMetSerArgLeuLys 343
 Db 726 GCTTCAAAAGATGTCATGCCCTTTGATTGCTTAAAGAGAGCTTCTTGTGAGATTAG 785
 Qy 344 ThrLeuGlyIleGlnValThrLysValTyrrGluGluGluTyrSerTyrrIleProValGly 363
 Db 786 ACACCTCGAATCCGATACTAAAGACTTACGAAGAGGATGCTCTATATCCAGTAGGT 845
 Qy 364 GlySerLeuProAsnThrGluGlnLysAsnLeuAlaPheGlyAlaAlaSerMetVal 383
 Db 846 GGTTCCTTCCCAACACGGAACAAAGAAATCTCGCCTTTGGCGTGCAGCTAGCATGTA 905

QY 384 HisProAlaThr 387
|||||
Db 906 CATCCGCAACA 917
RESULT 15
AAC44887
ID AAC44887 standard; DNA; 1039 BP.
XX AC AAC44887;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana DNA fragment SEQ ID NO: 44503.
XX KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0121180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0123788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134766.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.

Search completed: May 21, 2003, 22:28:53

reserSerCysProValTrp 16

Thu May '22 09:39:16 2003

us-09-701-395a-23.rng

Page 19

Job time : 396 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 21, 2003, 22:12:43 ; Search time 4278 Seconds
(without alignments)
3598.734 Million cell updates/sec

Title: US-09-701-395A-23

Perfect score: 2743

Sequence: 1 MELLGVRNLISCPVWTFGT.....VRHLLSDPSGAVVWVAYLER 529

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2.1/USPTO.spool/US09701395/runat_15052003.165750.3330/app_query.fasta_1.711
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09701395.ecgn.1.1.2586 @runat_15052003.165750.3330 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6 -FGAPEXT=7
-XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2743	100.0	1898	8	AF321536 Adonis pa
2	2728	99.5	1848	8	AF321535 Adonis pa
3	1966.5	71.7	1606	8	AY079371 Arabidops
4	1966.5	71.7	1836	8	AY040024 Arabidops
5	1959.5	71.4	1860	8	AR003634 Sequence
6	1959.5	71.4	1860	8	ATU50738 Arabodopsi
7	1934	70.5	1780	8	AF321538 Lactuca s
8	1934	70.5	1830	8	AF251016 Tagetes e
9	1934	70.5	1887	6	AR152414 Sequence
10	1913.5	69.8	1916	8	AY099485 Tagetes e
11	1909.5	69.6	1697	8	AY099485 Tagetes e
12	1892	69.0	2357	8	AF463497 Spinacia
13	1786	65.1	1550	8	AF486650 Citrus x
14	1633.5	59.6	3634	8	AF117257 Arabidops
15	1633.5	59.6	78423	8	AB024035 Arabidops
16	1518.5	55.4	1378	8	AF321537 Solanum t
17	1394	50.8	1272	6	AX350852 Sequence
18	1349	49.2	144879	8	AP003332 Oryza sat
19	1202	43.8	962	6	AX350851 Sequence
20	1113	40.6	-787	8	AF450280 Citrus si
21	874	31.9	605	8	AF229684 Daucus ca
22	783.5	28.6	2027	8	AF240787 Citrus si
23	778.5	28.4	1830	8	AF152246 Citrus x
24	778.5	28.4	3313	8	AY094582 Citrus si
25	775.5	28.3	2026	8	AF321534 Adonis pa
26	761	27.7	1785	8	X86221 C.annuum mR
27	761	27.7	1942	6	A58419 Sequence 1
28	755	27.5	1906	8	AF251017 Tagetes e
29	750.5	27.4	1650	6	AR022630 Sequence
30	750.5	27.3	1959	6	AR152412 Sequence
31	749.5	27.3	1650	8	LECRTL1
32	748	27.3	1537	8	AY091396 Arabidops
33	748	27.3	1946	8	ATU50739 Arabidops
34	748	27.3	1975	8	AY059749 Arabidops
35	748	27.3	2814	8	AF117256 Arabidops
36	748	27.3	87937	8	ATAC009400 Arabidops
37	747	27.2	1471	8	AY099484 Tagetes e
38	747	27.2	1661	8	ATHLYC
39	739	26.9	1590	6	AX350853 Sequence
40	737	26.9	1614	6	AR022629 Sequence
41	737	26.9	1614	8	NTCRTL1
42	736	26.8	497	8	AF212130 Daucus ca
43	729	26.6	1696	8	AF489520 Sanderson
44	719.5	26.2	1666	8	AF254793 Lycopersi
45	714.5	26.0	1498	8	STU272136 Solanum t

ALIGNMENTS

RESULT 1

```

AF321536
LOCUS       AF321536               1898 bp    mRNA    linear    PLN 28-FEB-2001
DEFINITION Adonis palaestina lycopene epsilon-cyclase mRNA, complete cds.
ACCESSION  AF321536
VERSION    AF321536.1 GI:12746308
KEYWORDS   .
SOURCE     Adonis palaestina.
            Adonis palaestina.
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales;
            Ranunculaceae; Adonis.
            1 (bases 1 to 1898).
            Cunningham, F. X. Jr. and Gantt, E.
            One ring or two? Determination of ring number in carotenoids by
            lycopene varepsilon-cyclases
            Proc. Natl. Acad. Sci. U.S.A. 98 (5), 2905-2910 (2001)
            21126984
            PUBMED 11226339
            2 (bases 1 to 1898)
            Cunningham, F. X. Jr.
            Direct Submission
            Submitted (16-NOV-2000) Cell Biology and Molecular Genetics,
            University of Maryland, Microbiology Building, Campus Drive,
            College Park, MD 20742, USA
            Location/Qualifiers
            1. .1898
            /organism="Adonis palaestina"
            /db_xref="taxon:151078"
            /tissue_type="Immature and developing flower buds"
            113. .1702
            /function="converts lycopene to delta-carotene and
            epsilon-carotene"
            /note="evidence by complementation in
            lycopene-accumulating Escherichia coli"
            /codon_start=1
            /evidence=experimental
            /product="lycopene epsilon-cyclase"
            /protein_id="AAK07432.1"
            /db_xref="GI:12746309"
            /translation="MELGVRLNLSGCPVMTFGFRLNLSKSLAYNIHRYGSSCRVDFQ
            VRADGSSRSVAYKEGFVDEDFIKAGSELFLVQMQTKSMERKQAKLADLPPI
            FGEVMDLVIIGCGPAGLSLAELGLKVLGIDPLPTNNYGNWEDFKDLGLER
            CIEHAKDTVYIENDAPVLIGRAIVRSRHLHEELKRCVSGSYLDSKVERITE
            AGDSHLVCNVEIFICRLATVAGSAGSKLLEYEYGVGPRVCVQAYGVEVEVNNP
            YDNLNMFMDRTYMQQLQCESEETFTFLYVMPSPTRLFEEETCLASDKAMPFDLL
            RKLSMLKLTIGIQVTKYEEWSEYIPVGSGLNPTNOKNLAFGAASAMVHPATGYSVV
            RSLSEAPKYASVIAKILKQNSAYVVSQSSAVNIQMAWSSLPKPKRQRAFFLVG
            LELIVQIDIEATRTFFTRFLPTMMWVGLGSLSLSDFLVLFMTMFVLAPNSMRMS
            LVRHLSDPSGVMVRAVLER"
            BASE COUNT  523 a  342 c  437 g  596 t
            ORIGIN

Alignment Scores:
Pred. No.:      8,25e-226      Length:      1898
Score:          2743.00      Matches:      529
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels:      0
DB:              8      Gaps:      0

US-09-701-395A-23 (1-529) x AF321536 (1-1898)

QY 1 MetGluLeuGlyValArgAsnLeuIleSerSerCysProValTrpThrPheGlyThr 20
DB 113 ATGGAACACTGTTGGTTCACACCTCATCTCTTCGTGCCCTGTGGACTTTTGAACA 172
QY 21 ArgAsnLeuSerSerIleValAlaTyrAsnIleHisArgTyrGlySerSerCysArg 40
DB 173 AGAAACCTTAGTGTTCACAACTAGCTTATACATACATACATGATATGTTCTTCTG 232
QY 41 ValAspPheGlnValArgAlaAspGlyGlySerGlySerArgSerValAlaTyrIys 60
DB 233 GTAGATTTTCAAGTAGAGCTGTATGTGGAAGCGGAGTAGAAGTCTCTGTTCTTAA 292

```

```

QY 61 GluGlyPheValAspGluGluAspPheIleIleGlyAlaGlySerGluLeuLeuPheVal 80
DB 293 GAGGGTTTGTGGATGAAGAGGATTTTATCAAGCTGGTGGTTCTGAGCTTTGTGTC 352
QY 81 GlnMetGlnGlnThrLysSerMetGluLysGlnAlaLysLeuAlaAspLysLeuProPro 100
DB 353 CAAATGCAGCAACAAAGTCTATGAGAACAGCCAGCTGCCGATAGTTGCCACA 412
QY 101 IleProPheGlyCysSerValMetAspLeuValIleGlyCysGlyProAlaGlyLeu 120
DB 413 ATACCTTTGGAGAAATCCGTGATGGACTTGGTTGTAATAGTTGTGACCTGCTGCTT 472
QY 121 SerLeuAlaAlaGluAlaAlaLysLeuGlyLeuLysValGlyLeuIleGlyProAspLeu 140
DB 473 TCACCTGGCTGCAGAGCTGCTAAGCTAGGTTGAAGTTGGCTTATGGTCTGATCTT 532
QY 141 ProPheThrAsnAsnTyrGlyValTrpGluAspGluPheLysAspLeuGluArg 160
DB 533 CTTTTTACAAATATATATGGTGTGTGGAAGACGAGTTCAAGAATCTTGGACTTGAACGT 592
QY 161 CysIleGluHisAlaTrpLysAspThrIleValTyrLeuAspAsnAspAlaProValLeu 180
DB 593 TGTATCGAGCATGCTTGAAGGACACCATCTATATCTTGTATATGATGCTCTGCTGCTT 652
QY 181 IleGlyArgAlaTyrGlyArgValSerArgHisLeuLeuHisGluLeuLeuLysArg 200
DB 653 ATGTGCTGTGATATGGACGAGTTAGTCGACATTTGCTACATGAGGAGTTCTGTAAGAG 712
QY 201 CysValGluSerGlyValSerTyrLeuAspSerLysValGluArgIleThrGluAlaGly 220
DB 713 TGTGTGGAGTCAGGTGATCATCTCTGGATTCTAAGTGAAGAGGATCACTGAAGCTGGT 772
QY 221 AspGlyHisSerLeuValCysGluAsnGluIlePheIleProCysArgLeuAlaThr 240
DB 773 GATGCCATAGCTTGTAGTTGTGAATAAGATCTTTATCCCTTCGAGGCTTGCTACT 832
QY 241 ValAlaSerGlyAlaAlaSerGlyLysLeuLeuGluTyrGluValGlyGlyProArgVal 260
DB 833 GTTCATCTCGAGCAGCTTCAGGGAACACTTTGGAGTATGAAGTAGTGCCCTCGTGT 892
QY 261 CysValGlnThrAlaTyrGlyValGluValGluAsnAsnProTyrAspProAsn 280
DB 893 TGTCTCAAAACCCCTTATGGGGTGGAGTTGAGGTGGAGAACCAATCCATACGATCCCAAC 952
QY 281 LeuMetValPheMetAspTyrAlaGlySerMetGlnLysLeuGlnCysSerGluGlu 300
DB 953 TTAATGTATTTCATGACTACAGAGACTATATCAACAGAAATACAGTCTCGGAGAA 1012
QY 301 GluTyrProThrPheLeuTyrValMetProMetSerProThrArgLeuPheGluGlu 320
DB 1013 GAATATCCAACATTTCTCTATGTCATGCCATGTGCCCAACAGACTTTTTTTTGGAGAA 1072
QY 321 ThrCysLeuAlaSerLysAspAlaMetProPheAspLeuLeuLysArgLysLeuMetSer 340
DB 1073 ACCTGTTTGGCTCAAAAGATGCCATGCCATTCGATCTACTCAAGAGAAACTGATGCA 1132
QY 341 ArgLeuLysThrLeuGlyIleGlnValThrLysValTyrGluGluTrpSerTyrIle 360
DB 1133 CGATTGAGACTCTGGTATCCAAAGTTTATCAAAAGTTTATGAAGAGATGGTCATATATT 1192
QY 361 ProValGlyGlySerLeuProAsnThrGluGlnLysAsnLeuAlaPheGlyAlaAla 380
DB 1193 CTTGTTGGTGTCTTTTACCAACACAGAGCAAAAGAACCTAGCATTTGGTGTGACGA 1252
QY 381 SerMetValHisProAlaThrGlyTyrSerValValArgSerLeuSerGluAlaProLys 400
DB 1253 AGCATGGTGCATCCAGCAACAGGCTATTTCGGTTGTACGGTCCTCAGAGAGCTCCAAA 1312
QY 401 TyrAlaSerValIleAlaLysIleLeuLysGlnAspAsnSerAlaTyrValValSerGly 420
DB 1313 TATGCTTCTGTAAATGCAAGATTTTGAAGCAAGATAACTCTGCGGTATGTGGTTCTTGA 1372

```

```

QY 421 GlnSerSerAlaValAsnIleSerMetGlnAlaTrpSerSerLeuTrpProLysGluArg 440
Db 1373 CAAAGTAGTCAGTAACATTTCAATGCAAGCAGTCTTTGGCCAAAGGACGGA 1432
QY 441 LysArgGlnArgAlaPhePheLeuPheGlyLeuGluLeuValGlnLeuAspIleGlu 460
Db 1433 AAACGTCAAGAGCATCTTCTTTTGGATTAGAGCTTATTGTGCAGCTAGATATGAA 1492
QY 461 AlaThrArgThrPhePheArgThrPhePheArgLeuProThrTrpMetTrpTrpGlyPhe 480
Db 1493 GCAACACAGAACATCTTTAGAACCTTCTCCGCTTCCCACTTGGATGGTGGGGTTTC 1552
QY 481 LeuGlySerSerLeuSerSerPheAspLeuValLeuPheSerMetTyrMetPheValLeu 500
Db 1553 CTTGGCTCTTCACTATCATCTTTCGATCTCGTCTTGTTCCTCCATGTACATGTTTGTG 1612
QY 501 AlaProAsnSerMetArgMetSerLeuValArgHisLeuLeuSerSerProSerGlyAla 520
Db 1613 GCGCCAAACAGCATGAGGATGTCACCTGTGAGACATTTGCTTTCAGATCCTTCTGTGCA 1672
QY 521 ValMetValArgAlaTyrLeuGluArg 529
Db 1673 GTTATGGTAAGACTTACCTCGAAAGG 1699

RESULT 2
AF321535 1848 bp mRNA linear PLN 28-FEB-2001
LOCUS
DEFINITION
ACCESSION AF321535
VERSION AF321535.1 GI:12746306
KEYWORDS
SOURCE
ORGANISM
Adonis palaestina.
Adonis palaestina.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales;
Ranunculaceae; Adonis.
REFERENCE
1 (bases 1 to 1848)
Cunningham,F.X. Jr. and Gantt,E.
One ring or two? Determination of ring number in carotenoids by
lycopene varepsilon-cyclases
Proc. Natl. Acad. Sci. U.S.A. 98 (5), 2905-2910 (2001)
21126984
JOURNAL
MEDLINE
PUBMED
1226339
REFERENCE
2 (bases 1 to 1848)
Cunningham,F.X. Jr.
Direct Submission
JOURNAL
TITLES
Submitted (16-NOV-2000) Cell Biology and Molecular Genetics,
University of Maryland, Microbiology Building, Campus Drive,
College Park, MD 20742, USA
FEATURES
Location/Qualifiers
1..1848
/organism="Adonis palaestina"
/db_xref="taxon:151078"
/tissue_type="Immature and developing flower buds"
116..1705
/function="converts lycopene to delta-carotene and
epsilon-carotene"
/notes="evidence by complementation in
lycopene-accumulating Escherichia coli"
/codon_start=1
/evidence-experimental
/product="lycopene epsilon-cyclase"
/protein_id="AAK07431.1"
/db_xref="GI:12746307"
/translation="MELIGVRNLISSCPWTGTRNLSSKLAYNHRYGSSCRVDFQ
VRADGSGRSVAKGEFVDEDFIKAGSSELFVQMOTKSMERQAKLADKLPIP
FGESYMDLWVTCGAPGLSLAAEAALGLKVLIGDLPFTNNYGVWDEFKDLGLER
CTEHWKDTIVYLDNDAPVLIIGRAYGRVSRHLHEELKRCVSGSVYLSNKRVERITE
ADGHSLVVCEVDEIPKRLATVASGAASKLLEYEVGGPRVCVOTAYGVEVENNP
YDNLVMPVMDVRYMQKLOCSSEYPTFLVYMPMSPTLFEETCLASKDAMPDNL
KKLMSRLKTLGIQTKIYEESWYIPVGGSLPTEQKNLAFGAASWVHPATGTSVV
RSLSEAPYASVIAKILKODNSAYVVGSSAVNTSMQASWLSWPKRRQRQRAFLFG
LELIVOLDIEATRTFTFRFLPTWWMGFLGSLSSFDLVLFMSYMEVLAPNSMRMS

```

BASE COUNT 522 a 334 c 420 g 572 t
ORIGIN

LVRHLLSDPSGAYVMVKAYLER"

Alignment Scores: 1.54e-224 Length: 1848
Pred. No.: 2728.00 Matches: 524
Score: 2728.00 Conservative: 5
Percent Similarity: 100.00%
Best Local Similarity: 99.05% Mismatches: 0
Query Match: 99.45% Indels: 0
DB: 8 Gaps: 0

US-09-701-395A-23 (1-529) x AF321535 (1-1848)

```

QY 1 MetGluLeuLeuGlyValArgAsnLeuIleSerSerCysProValTrpThrPheGlyThr 20
Db 116 ATGGAACCTACTGTGTGTCGCAACCTCATCTCTTCTCCCTGCTGACGCTTTTGGAAACA 175
QY 21 ArgAsnLeuSerSerSerLysLeuAlaTyrAsnIleHisArgTyrGlySerSerCysArg 40
Db 176 AGAAACCTTAGTACTTCAAAACTAGCTTATAACATACATCATGATGTTCTTCTTGTTAGA 235
QY 41 ValAspPheGlnValArgAlaAspGlyClySerGlySerArgSerSerValAlaTyrLys 60
Db 236 GTAGATTTTCAAGTGAGGGCTGTAGTGGTGAAGCGGAGTAGAAGCTTCTGTGCTTATAAA 295
QY 61 GluGlyPheValAspGluGluAspPheIleLysAlaGlyGlySerGluLeuLeuPheVal 80
Db 296 GAGGGTTTGTGGACGAGGAGGATTTATCAAGCTGTGTGCTTGAGCTTTTGTGTTGTC 355
QY 81 GlnMetGlnGlnThrLysSerMetGluLysGlnAlaLysLeuAlaAspLysLeuProPro 100
Db 356 CAATGTCAGCAACAAAGTCTATGAGAAACAGCCAGCTCGCGATAAGTTGCCACCA 415
QY 101 IleProPheGlyGluSerValMetAspLeuValIleGlyCysGlyProAlaGlyLeu 120
Db 416 ATACCTTTTCGAGAATCTGTGATGACTTGTGTTGTAATAGGTGTGTGGACCTGCTGCT 475
QY 121 SerLeuAlaGluAlaLysLeuGlyLeuLysValGlyLeuIleGlyProAspLeu 140
Db 476 TCACCTGGCTGCAGAGCTGTAACTAGCTGAGCTTGAAGCTTATTTGGCTTCTGATGCT 535
QY 141 ProPheThrAsnAsnTyrGlyValTrpGluAspGluPheLysAspLeuGluArg 160
Db 536 CCTTTTACAAATAATTATGTTGTGGAGACAGAGTTCAAAGATCTTGGACTTGAACGT 595
QY 161 CysIleGluHisAlaTrpLysAspThrIleValTyrLeuAspAsnAspAlaProValLeu 180
Db 596 TGTATCGAGCATGCTTGGAGGACCATCATCTATATCTTGACAATGATGCTCTGCTCTT 655
QY 181 IleGlyArgAlaTyrGlyArgValSerArgHisLeuLeuHisGluLeuLeuLysArg 200
Db 656 ATTGGTCTGCATATGACGAGTTAGCCGCACTTTGTCGATGAAGAGTTGCTGAAAAGG 715
QY 201 CysValGluSerGlyValSerTyrLeuAspSerLysValGluArgIleThrGluAlaGly 220
Db 716 TGTCTCAGTCAGCTGTATCATATCTGAATTTCTAAAGTGAAGAGGATCCTTGAAGCTGGT 775
QY 221 AspGlyHisSerLeuValValCysGluAsnGluIlePheIleProCysArgLeuAlaThr 240
Db 776 GATGGCCATAGTCTTGTAGTTTGTGAACACGACATCTTTATCCCTTTCAGGCTTGTCTACT 835
QY 241 ValAlaSerGlyAlaAlaSerGlyLysLeuLeuGluTyrGluValGlyGlyProArgVal 260
Db 836 GTTGCATCTGGACGAGCTTTCAGGAAACTTTTGGAGTATGAAGTAGGTGGCCCTCGTGT 895
QY 261 CysValGlnThrAlaTyrGlyValGluValGluValGluAsnAsnProTyrAspProAsn 280
Db 896 TGTGTCCAAACTGCTTATGTTGTGGAGTTGAGGTGGAGAACCAATCCATACGATCCCAAC 955
QY 281 LeuMetValPheMetAspTyrArgAspTyrMetGlnGlnLysLeuGlnCysSerGluGlu 300
Db 956 TTAATGTTATTTATGGACTACAGAGACTATATATGCAACAGAAATTTACAGTGTCTCGAAGA 1015

```

QY 301 GluTyrProThrPheLeuTyrValMetProMetSerProThrArgLeuPhePheGluGlu 320
 Db 1016 GAATATCAACATCTCTATGTATGTCATGCCATGTCGCCAACAAGACATTTTTTTTGAGGAA 1075
 QY 321 ThrCysLeuAlaSerLysAspAlaMetProPheAspLeuLeuLysArgLysLeuMetSer 340
 Db 1076 ACCTGTTGGGCTCAAAAGATGCCATGCTTTCATCTACTAGAGAGAAAATAATGTCA 1135
 QY 341 ArgLeuLysThrLeuGlyIleGlnValThrLysValTyrGluGluTyrSerTyrIle 360
 Db 1136 CGATTGAAGACTCTGGTATCCCAAGTTACAAAATTTATGAAGGAATGCTTTATATT 1195
 QY 361 ProValGlyGlySerLeuProAsnThrGluGlnLysAsnLeuAlaPheGlyAlaAla 380
 Db 1196 CCTGTTGGGGTCTTTTACCAACACAGAGCAAAAGAACCTAGCTTTGGTGTGCAGCA 1255
 QY 381 SerMetValHisProAlaThrGlyTyrSerValValArgSerLeuSerGluAlaProLys 400
 Db 1256 AGCATGTGTGATCCAGCAACAGGCTATTCGGTTGTAGCATCTATCAGAAAGTCCAAA 1315
 QY 401 TyrAlaSerValIleAlaLysIleLeuLysGlnAspAsnSerAlaTyrValValSerGly 420
 Db 1316 TATGCTTCTGTAATGCAAGATTTTGAAGCAAGATAACTCTGCATATGTTGTTCTGGA 1375
 QY 421 GlnSerSerAlaValAsnIleSerMetGlnAlaTrpSerSerLeuTrpProLysGluArg 440
 Db 1376 CAAGCAGTGCATGTAACATTTCAATGCAAGCATGAGCAGTCTTTGGCCAAAGAGCGA 1435
 QY 441 LysArgGlnArgAlaPhePheLeuPheGlyLeuGluLeuIleValGlnLeuAspIleGlu 460
 Db 1436 AAACGTCAAGAGCATCTTCTTTTCGGGTAGACCTTATGTGCGAGTATGATTTGAA 1495
 QY 461 AlaThrArgThrPhePheArgThrPhePheArgLeuProThrTrpMetTrpTrpGlyPhe 480
 Db 1496 GCACACAGACGTTCTTTAGAACCTTCTCCGCTTCCCAACTGGATGGTGGGTTC 1555
 QY 481 LeuGlySerSerLeuSerSerPheAspLeuValLeuPheSerMetTyrMetPheValLeu 500
 Db 1556 CTGGGTCTTCACTATCATCTTCGATCTTGTATGTTTCCATGATGATGTTTGTGTTG 1615
 QY 501 AlaProAsnSerMetArgMetSerLeuValArgHisLeuLeuSerAspProSerGlyAla 520
 Db 1616 GCCCGAAGACATGAGGATGTCATCTGTGACATTTGCTTTCAGATCCTTCTGTGCA 1675
 QY 521 ValMetValArgAlaTyrLeuGluArg 529
 Db 1676 GTTATGGTTAAGCTTACTCTGAAAAGG 1702
 RESULT 3
 LOCUS AY079371
 DEFINITION Arabidopsis thaliana putative lycopene epsilon cyclase (At5g57030)
 VERSION AY079371
 KEYWORDS Arabidopsis thaliana
 SOURCE Arabidopsis thaliana
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 1606)
 Yamada, K., Liu, S.X., Sakano, H., Pham, P.K., Banh, J., Etgu, P., Lee, J.M., Toriumi, M., Yu, G., Brooks, S., Chao, Q., Chen, H., Karlin-Neumann, G., Kim, C., Lam, B., Miranda, M., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.
 TITLE Arabidopsis Open Reading Frame (ORF) Clones
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1606)
 Yamada, K., Banh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L.,

Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinn, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.
 Direct Submission

TITLE
JOURNAL

Submitted (19-FEB-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA

COMMENT

The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL CDNAs (RAFL cDNA: RIKEN Arabidopsis Full-Length cDNA): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGECC (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RAFL CDNAs: Yamada, K., Banh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGECC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGECC) contributed equally to this work as PIs.

FEATURES

source

Location/Qualifiers

1..1606

/organism="Arabidopsis thaliana"

/db_xref="taxon:3702"

/chromosome="5"

/clone="U10203"

/note="This clone is in pUNI 51."

ecotype: Columbia"

1..1606

/gene="At5g57030"

1..1575

/gene="At5g57030"

/codon_start=1

/evidence="experimental"

/product="putative lycopene epsilon cyclase"

/protein_id="AAL85102.1"

/db_xref="GI:19310742"

/translation="MECVGARFAAAMAVSTFPSCRRKPPVVKRYSYNRIRGLCSV
 RASGGSSGSESCVAVRDEFADEDFKAGGSILFVQOQNDQKDEQSLVDKLPPI
 SIGDALDLVVICGPGAGLAAESAKLGLKVLGIDLPFTNNYGVWEDEFNDLGLQ
 KLEHWRETIYVDDDDKPTTIGRVRVSRRLLELLRRRCVSVYSSLRKVDISIT
 EASDGLRVACDDNNVPCRLATVSAAGASKLLOVEVGPVVCVQVAYGEVEVENS
 PYDPQMFMDYDYNKVRLEAEYPTFLYAMPTRKRLFEETCLASKDVMPFDL
 LKTKMLRLDTLIGIRLTKVEEWSYIPVGGSLPNTQKNLAFGAASAMVHATGYSV
 VRSLSAPKYASVIAELREETTQKINSINISQAWDTLWPPERKQRATFLGLALIV
 QPTEGIRSFERTFEFLPKMMQWGLGSLTSGDLVLFALYMFVISPNRLKGLNLH
 ISDPTGATMTKTLKV"

1576..1606

/gene="At5g57030"

439 a 300 c 391 g 476 t

BASE COUNT

ORIGIN

Alignment Scores:

Pred. No.: 2,6e-159 Length: 1606

Score: 1966.50 Matches: 377

Percent Similarity: 81.84% Conservative: 60

Best Local Similarity: 70.60% Mismatches: 80

Query Match: 71.69% Indels: 17

DB: 8 Gaps: 5

US-09-701-395A-23 (1-529) x AY079371 (1-1606)

QY 1 MetGluLeuGlyValArgAsnLeu-----IleSerSerCysProValTtp 16

Db 1 ATGGAGTGTGTGGGCTAGGAATTTTCGACGAATTCGGCGTTTCAACATTTCCGTCATGG 60

QY 17 ThrPheGlyThrArgAsnLeuSerSerLysLeuAlaTyrAsnIleHisArgTyrGly 36
 Db 61 AGTTGTCGAAGAAATTTCCAGTGGTTAAGAGATACAGCTATAGGAATATTCGTTTCGGT 120
 QY 37 SerSerCysArgValAspPheGlnValArgAlaAspGlyCly-----SerGlySerArg 54
 Db 121 ---TTCTGT-----AGTGTACAGCTAGGCGCGCGGAAGTTCCGGTAGCTAG 165
 QY 55 SerSerValAlaTyrLysGluGlyPheValAspGluGluAspPheIleLysAlaGlyGly 74
 Db 166 AGTTGTAGCGGTGAGAGAGATTTTCCTGACGAGAGAGATTTTGTGAAGCTGGTGGT 225
 QY 75 SerGluLeuPheValGlnMetGlnGlnThrLysSerMetGluLysGlnAlaLysLeu 94
 Db 226 TCTGAGATCTATTCTTCAATTCAGCAGCAACAAGATATGATGAACAGTCTAAGCTT 285
 QY 95 AlaAspLysLeuProIleProPheGlyGluSerValMetAspLeuValValIleGly 114
 Db 286 GTTGATAGTTCCTCTATATCAATTTGGTGTAGTGGTCTTTGGATCTAGTGGTATTTGGT 345
 QY 115 CysGlyProAlaGlyLeuSerLeuAlaAlaGluAlaLysLeuGlyLeuLysValGly 134
 Db 346 TGTGGTCTCTGTTAGCTTGGCTGCAGAAATCAGCTAAGCTTGGATTAAAGTTGGA 405
 QY 135 LeuIleGlyProAspLeuProPheThrAsnAsnTyrGlyValTrpGluAspGluPheLys 154
 Db 406 CTCATTGGTCCAGATCTCTCTTTACTAACAATTAGCGTGTGGGAAGATGAATCAAT 465
 QY 155 AspLeuGlyLeuGluArgCysIleGluHisAlaTrpLysAspThrIleValTrpLeuAsp 174
 Db 466 GATCTTGGCTCGCAAAATGTATGACATGTTGGAGAGAGACTATGTGTATCTGGAT 525
 QY 175 AsnAspAlaProValLeuIleGlyArgAlaTyrGlyArgValSerArgHisLeuLeuHis 194
 Db 526 GATGACAAAGCTATTACCATTTGSCCGCTTATGGAAGACTTAGTCGACGTTTGCCTCAT 585
 QY 195 GluGluLeuLysArgCysValGluSerGlyValSerTyrLeuAspSerLysValGlu 214
 Db 586 GAGGAGCTTTTGGAGAGGTGTGTCGAGTCTGAGTGTCTGCTACCTAGCTCGAAAGTTGAC 645
 QY 215 ArgIleThrGluAlaGlyAspGlyHisSerLeuValValCysGluAsnGluIlePheIle 234
 Db 646 AGCATAAACAGAGCTTCTGATGCGCTTAGACTTGTGTGTGACCAACAATAACGTCAT 705
 QY 235 ProCysArgLeuAlaThrValAlaSerGlyValAlaAlaSerGlyLysLeuLeuGluTrpGlu 254
 Db 706 CCTGCGAGCTTCCCACTGTCTGCTTGGAGCAGCTTCGGGAAGCTCTTGCAATAGCAA 765
 QY 255 ValGlyGlyProArgValCysValGlnThrAlaTyrGlyValGluValGluValGluAsn 274
 Db 766 GTTGTGGACCTAGAGTCTGTGTGCAAACTGCATAGCGGTGAGGTTGAGGTGGAAAT 825
 QY 275 AsnProTyrAspProAsnLeuMetValPheMetAspTyrArgAspTyrMetGlnGlnLys 294
 Db 826 AGTCCATATGATCCAGATCAATAGTTTTCATGATGATACAGATATATCTAAGAGAAA 885
 QY 295 LeuGlnCysSerGluGluTrpProThrPheLeuTyrValMetProMetSerProThr 314
 Db 886 GTTCCGAGCTTAGAAGCTGAGTATCCAACTTCTGCTGACCCATGCTATGACAAAGTCA 945
 QY 315 ArgLeuPhePheGluGluThrCysLeuAlaSerLysAspAlaMetProPheAspLeuLeu 334
 Db 946 AGACTCTTCTCGAGAGACATGTTTGGCTCAAAAGATGTCATGCCCTTTGATTTGCTA 1005
 QY 335 LysArgLysLeuMetSerArgLeuLysThrLeuGlyIleGlnValThrLysValTrpGlu 354
 Db 1006 AAAACCAAGCTCATGTATAGATTAGATACACTCGGAATTCGAATTTCTAAGACTTAGCAA 1065
 QY 355 GluGluTrpSerTyrIleProValGlyGlySerLeuProAsnThrGluGlnLysAsnLeu 374
 Db 1066 GAGGAGTGGTCCCTATATCCAGTTGTTGCTTCTGCAAAACACCGCAACAAGAAATCTC 1125
 QY 375 AlaPheGlyAlaAlaAlaSerMetValHisProAlaThrGlyTyrSerValValArgSer 394

Db 1126 GCCTTTGGTGTGTCGCGTAGCATCCCGCAACAGGCTATTAGTTGTGATCT 1185
 QY 395 LeuSerGluAlaProLysTyrAlaSerValIleAlaLysIleLeuLysGlnAspAsnSer 414
 Db 1186 TTGTCTGAAGCTCAAAATATGATCATGTCATCGCAGAGACTAAGAGAGAGACTACC 1245
 QY 415 AlaTyrValValSerGlyGlnSerSerAlaValAsnIleSerMetGlnAlaTrpSerSer 434
 Db 1246 AACAGATCAACAGT-----ATATTTCAAGACAAGCTTGGGATACT 1287
 QY 435 LeuTrpProLysGluArgLysArgGlnAlaPhePheLeuPheGlyLeuGluLeuIle 454
 Db 1288 TTATGGCCACCAGAAAGGAAAAAGACAGAGACATCTTCTCTTTGGTCTTGCACATCATA 1347
 QY 455 ValGlnLeuAspIleGluAlaThrArgThrPhePheArgThrPhePheArgLeuProThr 474
 Db 1348 GTTCAATTCGATACCGAAGGATTAGAAAGCTTCTCCGACTTCTCCGCCCTTCCAAA 1407
 QY 475 TrpMetTrpTrpGlyPheLeuGlySerSerLeuSerSerPheAspLeuValLeuPheSer 494
 Db 1408 TGGATGTGGCAAGGTTTCTAGGATCAACATTAACATCAGGAGATCTCGTTCTTTGCT 1467
 QY 495 MetTyrMetPheValLeuAlaProAsnSerMetArgMetSerLeuValArgHisLeuLeu 514
 Db 1468 TTATACATGTTTCGTCATTTCCACCAACAATTTGAGAAAAGTCTCATCAATCATCTCATC 1527
 QY 515 SerAspProSerGlyAlaValMetValArgAlaTyrLeuGlu 528
 Db 1528 TCTGATCAACCGGAGCAACCATGATATAAAACCTATCTCAAA 1569
 RESULT 4
 AY040024 1836 bp mRNA linear PLN 24-APR-2002
 LOCUS Arabidopsis thaliana putative lycopene epsilon cyclase (At5g57030)
 DEFINITION mRNA, complete cds.
 ACCESSION AY040024
 VERSION AY040024.1 GI:14532799
 KEYWORDS FLI-CDNA.
 SOURCE Arabidopsis thaliana.
 ORGANISM Arabidopsis thaliana.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsia.
 REFERENCE 1 (bases 1 to 1836)
 AUTHORS Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Chung,M.K.,
 Goldsmith,A.D., Lee,J.M., Quach,H.L., Toriumi,M., Yu,G., Bowser,L.,
 Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T.,
 Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B.,
 Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T.,
 Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K.,
 Davis,R.W., Ecker,J.R. and Theologis,A.
 Arabidopsis Full Length cDNA Clones
 TITLE Arabidopsis Full Length cDNA Clones
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1836)
 AUTHORS Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Chung,M.K.,
 Dale,J.M., Gibson,H.A., Goldsmith,A.D., Jiang,P.X., Lee,J.M.,
 Quach,H.L., Tang,C.C., Toriumi,M., Yu,G., Bowser,L., Carninci,P.,
 Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A.,
 Karlin-Neumann,G., Kawai,J., Kim,C., Koesema,E., Lam,B., Lin,J.,
 Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J.,
 Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Tracy,S.E.,
 Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.
 TITLE Direct Submission
 JOURNAL Submitted (11-JUN-2001) Plant Gene Expression Center, 800 Buchanan
 Street, Albany, CA 94710, USA
 COMMENT RIKEN Genomic Sciences Center (GSC) members carried out the
 collection and clustering of RAFI cDNAs (RAFL cDNA : 'RIKEN
 Arabidopsis Full-Length cDNA') : Seki,M., Narusaka,M., Ishida,J.,
 Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
 Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGECC (SSP) Consortium members carried out the

sequencing and annotation of the RAFL cDNAs: Yamada, K., Liu, S.X., Sakano, H., Pham, P.K., Banh, J., Chung, M.K., Dale, J.M., Gibson, H.A., Goldsmith, A.D., Jiang, P.X., Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Yu, G., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Koesema, E., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shinn, P., Southwick, A., Tracy, S.E., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC) contributed equally to this work as PIs.

FEATURES

source

Location/Qualifiers
1. .1836
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="5"

/clone="RAFL05-09-E07 (R10203)"
/note="This clone is in a modified pBluescript vector (Lambda ZAP) as a XhoI/SstI insert."
ecotype: Columbia"

1. .1836

/gene="At5g57030"

1. .105

/gene="At5g57030"

1. .2

/gene="At5g57030"

/note="compared to genomic sequence"

/replace="aa"

106. .1680

/gene="At5g57030"

/codon_start=1

/evidence="experimental"

/product="putative lycopene epsilon cyclase"

/protein_id="AAK64181.1"

/db_xref="GI:14532800"

/translation="MECVGARNFAAMVSTPSSCRKFFPVKRYSYRNFRLGCSV
RASGSGSECVAREDFADAEDEFVARGSEILFVQMNKMDQSKLVKLPPI
SIGDGLDVLGVCAGPALAAEASAKGLKVLGLPDLPTNNYGVDEFENDLQ
KCIHWRETIVYLDLDDKPIITGRAYGRVSRLLHEELLRSVGSVYLSKVDSSIT
EASDGLRVACDDNNVPCRLATVAGSAGKLLQYEVGGPRVCQVTAIGVEVENS
PYDDQWFMVYRDYTNKVRSLAEAYPTFLYAMPKSRLLFEETCLASKDVMPFDL
LKTLMRLDLTGLRILKTYEENSIFPVGSLPNTQNLAFGAASVHPATGYSV
VRLSEAPKVASVIAEILREETIKQINSIRQAWDLWPPERKQRAFLFLGLALIV
QPDGIRSFRTFRLPKWMMQGLSTLTSGLDVLFLALYMFVISPNNLRKGLINHL
ISDPGATMIKTYLKV"

1681. .1836

/gene="At5g57030"

BASE COUNT 530 a 324 c 440 g 542 t

ORIGIN

Alignment Scores:

Pred. No.: 3.1e-159

Score: 1966.50

Matches: 1836

Percent Similarity: 81.84%

Conservative: 60

Best Local Similarity: 70.60%

Mismatches: 80

Query Match: 71.69%

Indels: 17

Gaps: 5

US-09-701-395a-23 (1-529) x AY040024 (1-1836)

Qy 1 MetGluLeuLeuGluValArgAspLeu-----lleSerSerCysProValTrp 16

Db 106 ATGGAGTGTGTGGGTAGGAATTCGACGAATGCGGTTTCAACATTCCTCATGG 165

Qy 17 ThrPheGlyThrArgAsnLeuSerSerSerLysLeuAlaTrpAsnIleHisArgTyrGly 36

Db 166 AGTTGTGCAAGGAATTCAGTGGTAAAGATACAGCTATAGGAATATTCCTTCGGT 225

Qy 37 SerSerCysArgValAspPheGlnValArgAlaAspGlyGly-----SerGlySerArg 54

Db 226 ---TTGTGT-----AGTGTGAGAGCTAGCGGCGGCGGAGATTCCGGTAGTGAG 270

Qy 55 SerSerValAlaTrpLysGluGlyPheValAspGluGluAspPheIleLysAlaGlyGly 74

Db 55 SerSerValAlaTrpLysGluGlyPheValAspGluGluAspPheIleLysAlaGlyGly 74

Qy 55 SerSerValAlaTrpLysGluGlyPheValAspGluGluAspPheIleLysAlaGlyGly 74

Db 55 SerSerValAlaTrpLysGluGlyPheValAspGluGluAspPheIleLysAlaGlyGly 74

Qy 55 SerSerValAlaTrpLysGluGlyPheValAspGluGluAspPheIleLysAlaGlyGly 74

Db 55 SerSerValAlaTrpLysGluGlyPheValAspGluGluAspPheIleLysAlaGlyGly 74

Qy 55 SerSerValAlaTrpLysGluGlyPheValAspGluGluAspPheIleLysAlaGlyGly 74

Db 55 SerSerValAlaTrpLysGluGlyPheValAspGluGluAspPheIleLysAlaGlyGly 74

Db 271 AGTTGTGTAGCGGTGAGAGAAGATTTCGCTGACGAAGAAGATTTCGTGAAAGCTGGTGGT 330

Qy 75 SerGluLeuLeuPheValGlnMetGlnThrLysSerMetGluLysGlnAlaLysLeu 94

Db 331 TCTCAGATTCTATTGTTCAATCCAGACACAGATATGATGATGATGATGATGATGATGAT 390

Qy 95 AlaAspLysLeuProPheGluSerValMetAspLeuValIleGly 114

Db 391 GTTGATAGATTGCTCTATATCAATGCTGATGCTGCTTGGATCTAGTGGTATTGGT 450

Qy 115 CysGlyProAlaGlyLeuSerLeuAlaGluAlaLysLeuGlyValGly 134

Db 451 TGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 510

Qy 135 LeuIleGlyProAspLeuProPheThrAsnAspTyrGlyValTrpGluAspGluPheLys 154

Db 511 CTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 570

Qy 155 AspLeuGlyLeuGluArgCysIleGluHisAlaTrpLysAspThrIleValTrpLeuAsp 174

Db 571 GATCTTGGGCTGCAAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 630

Qy 175 AsnAspAlaProValLeuIleGlyArgAlaTrpGlyValSerArgHisLeuLeuHis 194

Db 631 GATGACAGCTTATACCATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 690

Qy 195 GluGluLeuLeuLysArgCysValGluSerGlyValSerLysSerLysValGlu 214

Db 691 GAGGAGCTTTTGGAGGAGTGTGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 750

Qy 215 ArgIleThrGluAlaGlyAspGlyHisSerLeuValValCysGluAsnGluPheIle 234

Db 751 ASCATAACAGAGCTTCTGATGCTTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 810

Qy 235 ProCysArgLeuAlaThrValAlaSerGlyAlaAlaSerGlyLysLeuGluTrpGlu 254

Db 811 CCCTGAGCTTGGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 870

Qy 255 ValGlyGlyProArgValCysValGlnThrAlaTrpGlyValGluValGluValGlu 274

Db 871 GTTGTGGACCTAGAGTGTGCTGCAAACTGCATACGCGTGGAGGTGAGGTGAGGTGAG 930

Qy 275 AsnProTrpAspProAsnLeuMetValPheMetAspTrpArgAspTyrMetGlnGly 294

Db 931 AGTCCATATGATCCAGATCAATGCTTTCATGCTTACAGATTTATCTAAGAGAGAA 990

Qy 295 LeuGlnCysSerGluGluTrpProThrPheLeuTrpValMetProMetSerProThr 314

Db 991 GTTCGAGCTTAGAAGCTGAGTATCCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1050

Qy 315 ArgLeuPheGluGluThrCysLeuAlaSerLysAspAlaMetProPheAspLeuLeu 334

Db 1051 AGACTCTTCTCGAGGAGAGATGTTGGCTCAAAAGATGTCATGCTGCTGCTGCTGCTGCT 1110

Qy 335 LysArgLysLeuMetSerArgLeuLysThrLeuGlyIleGlnValThrLysValTrpGlu 354

Db 1111 AAAACGAAGCTCATGTAGATTAGATACACTCGAATTCGAATTCGAATTCGAATTCGA 1170

Qy 355 GluGluTrpSerTrpIleProValGlySerLeuProAsnThrGluGlnLysAsnLeu 374

Db 1171 GAGGAGTGTCTATATCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1230

Qy 375 AlaPheGlyAlaAlaSerMetValHisProAlaThrGlyTyrSerValValArgSer 394

Db 1231 GCCTTTGGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1290

Qy 395 LeuSerGluAlaProLysTrpAlaSerValIleAlaLysIleLeuLysGlnAspAsnSer 414

Db 1291 TTGCTGAGAGCTCCAAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1350

Qy 415 AlaTrpValValSerGlyGlnSerSerAlaValAsnIleSerMetGlnAlaTrpSerSer 434

Db	1351	AAACAGATCAACAGT	-----AAATTTTCAAGACAAGCTTGGGACT	1392
QY	435	LeuTriProLysGluArgLysArgGlnArgAlaPhePheLeuPheGlyLeuGluLeu	454	
Db	1393	TTATGCCACACAGAAAGGAAACACAGAGAGCATCTTTCTCTTGGCTTGCACATCA	1452	
QY	455	ValGlnLeuAspIleGluAlaThrArgThrPhePheArgThrPhePheArgLeuProThr	474	
Db	1453	GTTCAATTCGATACCGAAGGCATTAGAAAGCTTCTCGTACTTCTTCGCGCTTCCAAA	1512	
QY	475	TriMetTriPrpGlyPheLeuGlySerSerLeuSerSerPheAspLeuValLeuPheSer	494	
Db	1513	TGGATGTGGCAAGGTTTCAGAGTCAACATTAACATCAGGAGATCTCGTCTCTTTGGCT	1572	
QY	495	MetTriMetPheValLeuAlaProAsnSerMetArgMetSerLeuValArgHisLeuLeu	514	
Db	1573	TTATACATGTTCTGTCATTTTCAACAAACAATTTGAGAAAAGGCTCTCATCAATCATCTCATC	1632	
QY	515	SerAspProSerGlyAlaValMetValArgAlaTyrLeuGlu	528	
Db	1633	TCTGATCCAAACCGGACCAACCATGATATAAAACCTATCTCAA	1674	
RESULT	5			
LOCUS	AR003634			
DEFINITION	Sequence 1 from patent US 5744341.	1860 bp	DNA	linear
ACCESSION	AR003634			
VERSION	AR003634.1	GI:3964893		
KEYWORDS	Unknown.			
SOURCE	Unknown.			
ORGANISM	Unclassified.			
REFERENCE	1 (bases 1 to 1860)			
AUTHORS	Cunningham,F.X. Jr. and Sun,Z.			
TITLE	Genes of carotenoid biosynthesis and metabolism and a system for screening for such genes			
JOURNAL	Patent: US 5744341-A 1 28-APR-1998;			
FEATURES	Location/Qualifiers			
source	1..1860			
BASE COUNT	531 a 327 c 448 g 554 t			
ORIGIN	/organism="unknown"			
Alignment Scores:				
Pred. No.:	1.26e-158	Length:	1860	
Score:	1959.50	Matches:	376	
Percent Similarity:	81.65%	Conservative:	60	
Best Local Similarity:	70.41%	Mismatches:	81	
Query Match:	71.44%	Indels:	17	
DB:	6	Gaps:	5	
US-09-701-395A-23 (1-529) x AR003634 (1-1860)				
QY	1	MetGluLeuLeuGlyValArgAsnLeu	-----IleSerSerCysProValTyr	16
Db	109	ATCGATGTGTGGGCTAGGAATTCGCCAGCAATCGCGGTTTCAACATTTCCGTCATGG	168	
QY	17	ThrPheGlyThrArgAsnLeuSerSerLysLeuAlaTyrAsnIleHisArgTyrGly	36	
Db	169	AGTTGTGCGAAGAAATTTCCAGTGGTTTACAGATACAGCTATAGGAATATTCGTTCCGGT	228	
QY	37	SerSerCysArgValAspPheGlnValArgAlaAspGly	-----SerGlySerArg	54
Db	229	---TTGTGT-----AGTGTCAGAGCTAGCGCGCGGAGAGTTCCCGTAGTGAG	273	
QY	55	SerSerValAlaTyrLysGluGlyPheValAspGluAspPheLysAlaGlyGly	74	
Db	274	AGTTGTGATCGCGTGAGAGAAGATTTTCGCTGACGACGAAGAGATTTTCTGAAAGCTGTGGT	333	
QY	75	SerGluLeuLeuPheValGlnMetGlnThrLysSerMetGluLysGlnAlaLysLeu	94	
Db	334	TCGAGATTCTATTGTTCATATGCAGCAGACAAAGATATGATGAACAGCTCTAAGCTT	393	

[illegible]

QY 157 GlyLeuGluArgCysIleGluHisAlaTrpLysAspThrIleValTyrLeuAspAsnAsp 176
 Db 570 GGACTTGGAGGCTGTATTAACATGTTGGGAGATCTAGTATATCTTGATGACAC 629
 QY 177 AlaProValLeuIleGlyArgAlaTyrGlyArgValSerArgHisLeuHisGluGlu 196
 Db 630 GATCCCATCTCTATAGTCTGCTATGACGAGTGTAGTCGTGATTTACTTACGAGGAG 689
 QY 197 LeuLeuLysArgCysValGluSerGlyValSerTyrLeuAspSerLysValGluArgIle 216
 Db 690 TTGTTGACTAGTGCATGAGTCAGCGTTTCATATCTGAGTCCAAAGTGAACGGATT 749
 QY 217 ThrGluAlaGlyAspGlyHisSerLeuValValCysGluAsnGluIlePheIleProCys 236
 Db 750 ACTGAAGCTCCAAATGGCTTAAGCTCATAGATGTGAAGCAATATCAATTCATCCATGC 809
 QY 237 ArgLeuAlaThrValAlaSerGlyAlaLaSerGlyLeuLeuGluValGluValGly 256
 Db 810 AGGCTTGCTACTGCTCTGAGCAGCTTCTGGAAACTTTTCAGTATGAACCTTGGC 869
 QY 257 GlyProArgValCysValGlnThrAlaTyrGlyValGluValGluValGluAsnPro 276
 Db 870 GTGTCCTGCTGTTGCTTCAACACGCTTATGATAGAGTTGAGTTGAAAGCATACCC 929
 QY 277 TyrAspProAsnLeuMetValPheMetAspTyrArgAspTyrMetGlnGlnLysLeuGln 296
 Db 930 TATGATCCAAAGCTTAATGCTTTCATGATATAGACTACACCAACATAAATCTCAA 989
 QY 297 CysSerGluGluGluTyrProThrPheLeuTyrValMetProMetSerProThrArgLeu 316
 Db 990 TCACCTAGAAGCACATATCCAACTTTTGTATGTCATGCCAATCTCTCCAACTAAAGTA 1049
 QY 317 PhePheGluGluThrCysLeuAlaSerLysAspAlaMetProPheAspLeuLysArg 336
 Db 1050 TTCCTTGAGGAACCTGTTGCTTCAAAAGAGGCGCATGCCCTTTGAGTTATGAAGACA 1109
 QY 337 LysLeuMetSerArgLeuLysThrLeuGlyIleGlnValThrLysValTyrGluGluGlu 356
 Db 1110 AAACCTCATGCAAGATTAAAGCATATGGGATCGGATTCGAATACCAAACTTATGAAGAGAA 1169
 QY 357 TrpSerTyrIleProValGlyGlySerLeuProAsnThrGluGlnLysAsnLeuAlaPhe 376
 Db 1170 TGGTCATATATCCAGTAGTGGATCTCTTACCAATACCGAGCAAAAGAACCTTGATTT 1229
 QY 377 GlyAlaAlaLaSerMetValHisProAlaThrGlyTyrSerValValArgSerLeuSer 396
 Db 1230 GGTGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 1289
 QY 397 GluAlaProLysTyrAlaSerValIleAlaLysIleLeuLysIleLeuLysIleLeuLys 416
 Db 1290 GAAGCTCCTAATATATGACAGCATTAATGCAAGATTTTAGGAAAGGAATTCAAACAG 1349
 QY 417 ValValSerGlyGlnSerAlaValAsnIleSerMetGlnAlaTrpSerSerLeuTrp 436
 Db 1350 ATGCTTGATCATGGAAGATACACAACTCTCAAGCAAGCTTGGAAACACTTTGG 1409
 QY 437 ProLysGluArgLysArgGlnArgAlaPhePheLeuPheGlyLeuGluLeuValGln 456
 Db 1410 CCCCTTGAAGGAAACACAGACAGACATCTTCTCTTGGATTTAGCAGTATGCTGAC 1469
 QY 457 LeuAspIleGluAlaThrArgThrPhePheArgThrPhePheArgLeuProThrTrpMet 476
 Db 1470 ATGGATATTAGGGGACCCGACATCTTCCGAGCTTCTCCGCTTGGCCACATGGATG 1529
 QY 477 TrpTrpGlyPheLeuGlySerLeuSerSerPheAspLeuValLeuPheSerMetTyr 496
 Db 1530 TGGTGGGGGTTCTTGGATCTTCTTATCATCACTACGATGATGATATTTGGCTTTTAC 1589
 QY 497 MetPheValLeuAlaProAsnSerMetArgMetSerLeuValArgHisIleLeuSerAsp 516
 Db 1590 ATGTTTATCATAGCAGCGCATGCTGAGATGGCTGCTGATGATGATGATGATGATGATG 1649
 QY 517 ProSerGlyAlaValMetValArgAlaTyrLeu 527

Db 1650 CCGACAGGAGGACAAATGTTAAAGCGTATCTC 1682
 RESULT 10
 AY099485
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

AY099485 1916 bp mRNA linear PLN 09-JUN-2002
 Tagetes erecta epsilon cyclase (lcy-e) mRNA, complete cds.
 AY099485.1 GI:21360358

African marigold.

ORGANISM

Tagetes erecta
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids II; Asterales; Asteraceae; Asteroideae;
 Heliantheae; Tagetes.

REFERENCE

AUTHORS

TITLE

1 (bases 1 to 1916)
 and Paredes-Lopez, O.
 Isolation and functional characterization of carotenogenic genes
 (psyl, pds, lcy-b and lcy-e) from marigold (Tagetes erecta) and
 ultrastructural changes during flowering development

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

gene

CDS

BASE COUNT

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-09-701-395A-23 (1-529) x AY099485 (1-1916)

QY 22 AsnLeuSerSerSerLysLeuAlaTyrAsnIleHisArgTyrGlySerSerCysArgVal 41

Db 191 CACATGACGCAACATGCGGCTTTTACATGCGCTTTTATGACTACATCATAC 250

QY 42 AspPheGlnValArgAlaAspGlyGlySerGlySerArgSerValAlaTyrLysGlu 61

Db 251 ACGAAGCAAAATTAAGTGC-----AACGCTGCTAAAGCCAGCTAGTCGTTAAACAA 301

BASE COUNT 580 a 361 c 427 g 548 t

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-09-701-395A-23 (1-529) x AY099485 (1-1916)

QY 22 AsnLeuSerSerSerLysLeuAlaTyrAsnIleHisArgTyrGlySerSerCysArgVal 41

Db 191 CACATGACGCAACATGCGGCTTTTACATGCGCTTTTATGACTACATCATAC 250

QY 42 AspPheGlnValArgAlaAspGlyGlySerGlySerArgSerValAlaTyrLysGlu 61

Db 251 ACGAAGCAAAATTAAGTGC-----AACGCTGCTAAAGCCAGCTAGTCGTTAAACAA 301

BASE COUNT 580 a 361 c 427 g 548 t

ORIGIN

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-09-701-395A-23 (1-529) x AY099485 (1-1916)

QY 22 AsnLeuSerSerSerLysLeuAlaTyrAsnIleHisArgTyrGlySerSerCysArgVal 41

Db 191 CACATGACGCAACATGCGGCTTTTACATGCGCTTTTATGACTACATCATAC 250

QY 42 AspPheGlnValArgAlaAspGlyGlySerGlySerArgSerValAlaTyrLysGlu 61

Db 251 ACGAAGCAAAATTAAGTGC-----AACGCTGCTAAAGCCAGCTAGTCGTTAAACAA 301

QY 62 GlyPheValAspGluAspPheIleLysAlaGlyGlySerGluLeuLeuPheValGln 81
 Db 302 GAGATTGAGGAGGAGAGATATGTAAGACCGGTGATCGGAGTGCTTTTGTTCAA 361
 QY 82 MetGlnGlnThrLysSerMetGluLysGlnAlaLysLeuAlaAspLysLeuProIle 101
 Db 362 ATGCAACAGAAATAGTCCATGATGACACAGCTAGGCTATCCCAAAAGCGTCCCAAGGTA 421
 QY 102 ProPheGly-----GluSerValMetAspLeuValIleGlyCysGly 116
 Db 422 CCATAGGAGGAGGAGGAGACAGTACTGTATACCTGGATTGGTGTAAATGGTGGT 481
 QY 117 ProAlaGlyLeuSerLeuAlaGluAlaLysLeuGlyLeuLysValGlyLeuIle 136
 Db 482 CTGCTGGCTGCTCTCTGCTGAGAAATCAGCCAAAGTAGGCTGAATGTCGCACTATC 541
 QY 137 GlyProAspLeuProPheThrAsnAsnTyrGlyValTrpGluAspGluPheLysAspLeu 156
 Db 542 GGCCTGTATCTCTTTTACAAATACTATGTTGGTGGGAGGATGAATTTATAGTCTT 601
 QY 157 GlyLeuGluArgCysIleGluHisAlaTrpLysAspThrIleValTyrLeuAspAsnAsp 176
 Db 602 GACTTGGAGGCTGATTAACATGTTGGGAGATGACTGTATGATATCTGTAGCAAC 661
 QY 177 AlaProValLeuIleGlyArgAlaTyrGlyArgValSerArgHisLeuLeuHisGluGlu 196
 Db 662 GATCCCATCTCATAGTCTGCTGATGACGAGTATGCTGATTTACTTACGAGGAG 721
 QY 197 LeuLeuLysArgCysValGluSerGlyValSerTyrLeuAspSerLysValGluArgIle 216
 Db 722 TTGTGACTAGTGTGCTGAGTGTGAGTGTGATCTGAGCTCCAAAGTGGAAAGGAT 781
 QY 217 ThrGluAlaGlyAspGlyHisSerLeuValValCysGluAsnGluIlePheIleProCys 236
 Db 782 ACTGAGCTCCAAATGGCTAAGTCTATGAGTGTGAGGCAATATCAAAATCTCATGC 841
 QY 237 ArgLeuAlaThrValAlaSerGlyAlaAlaSerGlyLysLeuLeuGluTrpGluValGly 256
 Db 842 AGGCTGTCTACTGCTGAGCAGCTCTGGGAAATCTTGGATGATGAATGAGTGGC 901
 QY 257 GlyProArgValCysValGlnThrAlaTyrGlyValGluValGluValGluAsnAsnPro 276
 Db 902 GTGCCCGGTGTTGGCTTCAACAGCTTATGTTAGAGGCTGAGTGTGAAGCATACCC 961
 QY 277 TyrAspProAsnLeuMetValPheMetAspTyrArgAspTyrMetGlnGlnLysLeuGln 296
 Db 962 TATGATCCAAAGCTATGTTTTCATGATATGAGACTACACCAAAACATAAATCTCAA 1021
 QY 297 CysSerGluGluGluTyrProThrPheLeuTyrValMetProMetSerProThrArgLeu 316
 Db 1022 TCACTAGAAGCAACAATATCCAACTTTTGTATGTCATGCAATGCTCCAACTAAAGTA 1081
 QY 317 PhePheGluGluThrCysLeuAlaSerLysAspAlaMetProPheAspLeuLeuLysArg 336
 Db 1082 TTCCTTGGAGAACTGTTGGCTTCAAAAGAGCGCATGCTTTGAGTTATTGAAGACA 1141
 QY 337 LysLeuMetSerArgLeuLysThrLeuGlyIleGlnValThrLysValTyrGlu----- 354
 Db 1142 AAACATGTCAGATTAAGACTATGAGGATGCGGATCCGAATACCAAACTATGAAAGATAT 1201
 QY 355 -----GluGluTrpSerTyrIleProValGlyGlySerLeuPro 367
 Db 1202 CTTGTTGCTTGTCAATATTTTGAAGAATGTCATATATTCACAGTAGTGGATCCCTCCA 1261
 QY 368 AsnThrGluGlnLysAsnLeuAlaPheGlyAlaAlaSerMetValHisProAlaThr 387
 Db 1262 AATACCGAGCAAAAGAACCTTCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1321
 QY 388 GlyTyrSerValValArgSerLeuSerGluAlaProLysTyrAlaSerValIleAlaLys 407
 Db 1322 GCATATTCGGTGTAGATCAGTGTGAGGAGCTCCCAATATTCAGCAGCAATATGCAAG 1381

QY 408 IleLeuLysGlnAspAsnSerAlaTyrValValSerGlyGlnSerSerAlaValAsnIle 427
 Db 1382 ATTTTAGGAAGGAATTCAAAACAGATGCTTGTATCTTGGAGATACACACCAACATC 1441
 QY 428 SerMetGlnAlaTrpSerSerLeuTrpProLysGluArgLysArgGlnArgAlaPhePhe 447
 Db 1442 TCAAAGCAAGCTTGGAAACACTTTGGCCCTTGAAGGAAAGACAGAGACATCTCTTT 1501
 QY 448 LeuPheGlyLeuGluLeuIleValGlnLeuAspIleGluAlaThrArgThrPhePheArg 467
 Db 1502 CTCCTTGGATTAGCAGTATGTCAGATGATATGAGGGAGCCCGCACATCTCTCCGG 1561
 QY 468 ThrPhePheArgLeuProThrTrpMetTrpGlyPheLeuGlySerSerLeuSerSer 487
 Db 1562 ACTTCTCCGCTTCCACATGATGTTGGTGGGGTTCCTGGATCTTCGTATCATCA 1621
 QY 488 PheAspLeuValLeuPheSerMetTyrMetPheValLeuAlaProAsnSerMetArgMet 507
 Db 1622 ACTGACTGTATAATATTGCGTATTTACATGTTTATCATGACCCGATAGCCTGAGAATG 1681
 QY 508 SerLeuValArgHisLeuLeuSerAspProSerGlyAlaValMetValArgAlaTyrLeu 527
 Db 1682 GGTCTGGTGTAGCATTTGCTTCTGACCCGACAGGAGCAATGTTAAAGCGTATCTC 1741
 RESULT 11
 LELYCOCYC
 LOCUS LELYCOCYC 1697 bp mRNA linear PLN 31-MAR-1998
 DEFINITION Lycopersicon. esculentum mRNA for lycopen epsilon-cyclase.
 ACCESSION Y14387
 VERSION Y14387.1 GI:3005982
 KEYWORDS CrtL-e-1 gene; lycopen epsilon-cyclase.
 SOURCE Lycopersicon esculentum.
 ORGANISM Lycopersicon esculentum.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
 Lycopersicon.
 REFERENCE 1 (bases 1 to 1697)
 Ronen, G., Cohen, M., Zamir, D. and Hirschberg, J.
 AUTHORS Regulation of expression of the gene for lycopen epsilon cyclase
 TITLE during fruit ripening of tomato
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 1697)
 Hirschberg, J.
 AUTHORS Direct Submission
 TITLE Submitted (31-JUL-1997) J. Hirschberg, The Hebrew University of
 JOURNAL Jerusalem, Dept of Genetics, Givat Ram, Jerusalem 91904, ISRAEL
 FEATURES
 location/Qualifiers
 1. 1697
 /organism="Lycopersicon esculentum"
 /cultivar="VF36"
 /db_xref="taxon:4081"
 /tissue_type="leaf"
 6. 1586
 /gene="CrtL-e-1"
 /gene="CrtL-e-1"
 /functions="converts lycopen to delta-carotene"
 /codon_start=1
 /product="lycopen epsilon-cyclase"
 /protein_id="CAA74745.1"
 /db_xref="GI:3005983"
 /db_xref="SPTREMBL:O65837"
 /translation="MECVGVQNVGMAVLTRPLRMWSGGELCOEKSTFLAYEYESK
 CNSSSSDSCVDEKEDFADEEDYIKAGGSQVLFVQMQKKDKMDQKSLDELROISAG
 QTVLDLVIGCGPAGLAAASAKLGLNVLGVLDPDLPFTNNNGVQWDEKDLQACI
 EHWYRDTIVYLDDEDTLIGRAYGRVSRHFLHEELLKRCVAGLYLNSKVDRIVEAT
 NGSLVECEGDVVPICRFTVAGSAAGSKFLQYELGSPRVSVQVAYGVEVDNPF
 PSLWFMVDRDYLHDAQSLKAVPTFLYAMPSPTRVFVEETCLASKDAMPFDLLKK
 KMLRLNTLGVRIKEIYEESVYIPVCGSLPTEQKTLAFCAASVHVHPATGYSVWS
 LSPAPKASVLANILRHQYHKNMLTSSSIFSIQTQANTLHPQERKQRFFFLGLAL
 ILQDLGEGIRSFRAFRVPKVMQWQGLSGSLSSLDLMLFAFTNFIAPDMRRGLIR
 HLLSDPTGATLIRTYTF"

BASE COUNT	473 a	308 c	403 g	513 t
ORIGIN				
Alignment Scores:				
Pred. No.:	2.16e-154	Length:	1697	
Score:	1909.50	Matches:	368	
Percent Similarity:	81.19%	Conservative:	68	
Best Local Similarity:	68.53%	Mismatches:	78	
Query Match:	69.61%	Indels:	23	
DB:	8	Gaps:	6	
US-09-701-395a-23 (1-529) x LELYCOCY (1-1697)				
QY	1	MetGluLeuLeuGlyValAlaArgAsnLeuLeuSerSerCysProVal	-----15	
Db	6	ATGGAGTGTGTGGAGTTCACAAATGTTGGAGCAATGCGAGTTCCTTAACCGCTCCGAGATTG	65	
QY	16	-----TrpThrPheGlyThrArgAsnLeuSerSerSerLys	-----LeuAlaTyr	30
Db	66	AACCGTGTGGTGGGAGGA-----GAGTTATGCCAAGAAAAGCATCTTTTGGCGGTAT	119	
QY	31	AsnIleHisArgTyrGlySerSerCysArgValAspPheGlnValArgAlaAspGlyGly	50	
Db	120	-----GAGCAGTATGAAGATAATGT-----AATAGCAGT	149	
QY	51	SerGlySerArgSerValAlaTyrLysGluGlyPheValAspGluGluAspPheIle	70	
Db	150	AGTGTAGTACACAGTGTGTAGTTGATAAGAGATTTTCTGATGAAGAAGATTATATA	209	
QY	71	LysAlaGlyGlySerGluLeuLeuPheValGlnMetGlnThrLysSerMetGluLys	90	
Db	210	AAAGCCGGTGGTGGCAACTGTATTGTTCAATCCAGCAGAAAAGATATGATGATCAG	269	
QY	91	GlnAlaLysLeuAlaAspLysLeuProPheGlyGluSerValMetAspLeu	110	
Db	270	CAGTCAAGCTTCTGATGAGTTACAGCAAAATCTCTGGCAACACCTACTGATTA	329	
QY	111	ValValIleGlyCysGlyProAlaGlyLeuSerLeuAlaAlaGluAlaLysLeuGly	130	
Db	330	GTGGTAATCGGTGTGGTCTGCTGTCTTCTTGGCGGAGTCAGCTAAATGGGG	389	
QY	131	LeuLysValGlyLeuLeuGlyProAspLeuProPheThrAsnAsnTyrGlyValTrpGlu	150	
Db	390	TTGAAGTGGGCTGTGGGCTGATCTCTCTTTCAACAACAACATATGGTGTGGAG	449	
QY	151	AspGluPheLysAspLeuGlyLeuGlyArgCysIleGluHisAlaTyrLysAspThrIle	170	
Db	450	GACGAGTTCAAGATCTTGGTCTTCAAGCCTGCATGAACATGTTGGCGGAGTACCATT	509	
QY	171	ValTyrLeuAspAsnAlaProValLeuLeuLeuGlyArgAlaTyrGlyArgValSerArg	190	
Db	510	GTATATCTGTATGATGATGAACCTATTCTTATGGCGTGCCTATGGAAGATTAGTCGC	569	
QY	191	HisLeuLeuHisGluGluLeuLysArgCysValGluSerGlyValSerTyrLeuAsp	210	
Db	570	CATTCTTCGCAGGAGTACTCAAAAGGTGTGTGGAGCAGGTGTTGTATCTAAAC	629	
QY	211	SerLysValGluArgIleThrGluAlaGlyAspGlyHisSerLeuValValCysGluAsn	230	
Db	630	TCGAAAGTGGATAGATTGTGGGCCACAAATGGCCAGAGTCTGTAGAGTGGCAGGT	689	
QY	231	GluIlePheIleProCysArgLeuAlaThrValAlaSerGlyAlaAlaSerGlyLysLeu	250	
Db	690	GATGTGTGATTCCTCAGGTTTGTACATGTTGTCATGGGGGAGCCTCGGGAAATTC	749	
QY	251	LeuGluTyrGluValGlyGlyProArgValCysValGlnThrAlaTyrGlyValGluVal	270	
Db	750	TTGCAGTATGAGTTGGGAGTCTCAGTTCTGTTCAACACCTATGGAGTGGAGTT	809	
QY	271	GluValGluAsnProTyrAspProAsnLeuMetValPheMetAspTyrArgAspTyr	290	
Db	810	GAGTTGATACAAATCCATTTGACCGAGCCTGATGTTTCTATGATATAGATTAT	869	

QY	291	MetGlnGlnLysLeuGlnCysSerGluGluGluTyrProThrPheLeuTyrValMetPro	310
Db	870	CTCAGACACGACGCTCAATCTTTAGAGCTAAATATATCAACATTTCTTTATGCATGCC	929
QY	311	MetSerProThrArgLeuPheGluGluThrCysLeuAlaSerLysAspAlaMetPro	330
Db	930	ATGCTCCACACAGAGTCTTTTTCGAGGAACCTGTTGGCTTCAAAAGATGCAATGCCA	989
QY	331	PheAspLeuLysArgLysLeuMetSerArgLeuLysThrLeuGlyIleGlnValThr	350
Db	990	TTGATCTGTAAAGAAAACCTGATGCTAGATTGAACACCTTGGTGAAGATATAA	1049
QY	351	LysValTyrGluGluThrPheTyrIleProValGlySerLeuProAsnThrGlu	370
Db	1050	GAATTTACGAGGAGGAATGGTCTTACATACGGTGGTGGATCTTTGCCAAATACAGAA	1109
QY	371	GlnLysAsnLeuAlaPheGlyAlaAlaLaSerMetValHisProAlaThrGlyTyrSer	390
Db	1110	CAAAAACACTTGTGCTGCTGCTAGCTGGTTCATCCAGCCACAGGTTATTCA	1169
QY	391	ValValArgSerLeuSerGluAlaProLysTyrAlaSerValIleAlaLysIleLeuLys	410
Db	1170	GTGCTGAGATCACTTCTGAAGCTCCAAATCGCCTCTGTACTTGCAAAATATATAGA	1229
QY	411	GlnAspAsnSerAlaTyrValValSerGlyGlnSerSerAlaValAsnIleSerMetGln	430
Db	1230	CAACATTATAGCAACACATGCTTACC--AGTTCAAGTATCCCGAGTATATCAACTCA	1286
QY	431	AlaTrpSerSerLeuTrpProLysGluArgLysArgGlnArgAlaPhePheLeuPheGly	450
Db	1287	GCTTGAACACTCTTTGGCCACAAAGAACGACAAAGATCGTTTTCTCTATTTGGA	1346
QY	451	LeuGluLeuIleValGlnLeuAspIleGluAlaThrArgThrPheArgThrPhePhe	470
Db	1347	CTGCTCGATGATTCGAGCTGGATTTGAGGGGATAGAGTTCATTTTCCGCGCATTTCT	1406
QY	471	ArgLeuProThrTrpMetTyrPheLeuGlySerSerLeuSerSerPheAspLeu	490
Db	1407	CGTGTGCAAAATGGATGGCAGGATTTCTTGGTTCAAGTCTTCTTTCAGCAGACCTC	1466
QY	491	ValLeuPheSerMetTyrMetPheValLeuAlaProAsnSerMetArgMetSerLeuVal	510
Db	1467	ATGTTATTTGCTTCTACATGTTTATTATTCACCAATGACATGACAAAGGCTTGATC	1526
QY	511	ArgHisLeuSerAspProSerGlyAlaValMetValArgAlaTyrLeu	527
Db	1527	AGATCTTTTATCTGATCTACTGTGTCACATTTGATAAGAACTTATCTT	1577

RESULT 12	AF463497	2357 bp	mRNA	linear	PLN 30-JAN-2002
LOCUS	Spinacia oleracea	lycopene epsilon-cyclase (lec) mRNA, complete			
DEFINITION	cds.				
ACCESSION	AF463497				
VERSION	AF463497.1	GI:18419660			
KEYWORDS	Spinacia oleracea.				
SOURCE	Spinacia oleracea				
ORGANISM	Spinacia oleracea				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.				
AUTHORS	DeSouza, M.L., Kollmann, S.R. and Schroeder, W.A.				
TITLE	Production of Lutelin in Microorganisms				
JOURNAL	2 (bases 1 to 2357)				
REFERENCE	DeSouza, M.L., Kollmann, S.R. and Schroeder, W.A.				
AUTHORS	Direct Submission				
TITLE	Submitted (27-DEC-2001) Central Research, Cargill Inc., P.O. Box				
JOURNAL	5702, Minneapolis, MN 55440, USA				
FEATURES	Location/Qualifiers				
source	1. 2357				
	/organism="Spinacia oleracea"				

```

/cultivar="Melody"
/db_xref="taxon:3562"
/tissue_type="actively-growing leaves"
1..2357
/gene="lec"
CDS
264..1817
/gene="lec"
/function="converts lycopene to delta-carotene"
/codon_start=1
/product="lycopene epsilon-cyclase"
/protein_id="AAL69394.1"
/db_xref="GI:18419661"
/translation="MVQRQNFQTFCEWPNSSVVVVESSRRSSSVLRSANSDDSC
VPADFNEEDFIKAGSELYVQMOQNKAMDCYSKIDKLRIQSDANELLDMVIG
CPAGLALAAESAKGLKVLGVPOLPTNNYGVWDEFRALGLGCIHVRDITIV
IDDDNPIYGRSVKVSROLLHKLVRCLSESVYLNKAKVENIMEGDPGRHVACER
GWTIPCLRVTVASGASCKLLEVEYGGPRVCVQTAYGVEVEVENSYPDNNVFMQYR
DYTKLSVGSLEAKYFTFLIAMPISPTRIFFETCLASVDAMPFDLLKKLMRLQTMG
VRITIKYEEWSYIPVGSPLNTEQRNLFAASNMVHPATGYSVVRSSEAPKYASA
IANLIKNDLSKNAILQRQSVGNISQAWNTLPQERKQRAFFLGLSLIVLDIEGI
RTFFRTFVRPKWMEGFLGSLSSADLILFAFYMEFFIAPNDLRMLIRHLLSDPTGA
TMIRTYITL"

```

BASE COUNT 654 a 469 c 546 g 688 t

Alignment Scores:

Pred. No.: 1.05e-152 Length: 2357
 Score: 1892.00 Matches: 355
 Percent Similarity: 83.70% Conservative: 61
 Best Local Similarity: 71.43% Mismatches: 75
 Query Match: 68.98% Indels: 6
 DB: 8 Gaps: 1

US-09-701-395a-23 (1-529) x AF463497 (1-2357)

```

QY 37 SerSerCysArgValAspPheGlnValArgAlaAspGlySerGlySerArgSerSer 56
Db 318 TCGAACCTGTAGTAGAATGTAGTAGTCGTAGGAGTAGTAGTGTTCGAGGAGTGGC 377
QY 57 -----ValAlaTyrLysGluGlyPheValAspGluGluAspPheIle 70
Db 378 AATAGCGACAGTAGTAGTCGTAATTCGCCAGAGAGATTTCCGACAGAGAGATTCATC 437
QY 71 LysAlaGlySerGluLeuLeuPheValGlnMetGlnGlnThrLysSerMetGluLys 90
Db 438 AAAGCTGGTGGTCCGAGCTCTTTATGTTCAAAATGAGCAGAGATTAAGCTATGGATTG 497
QY 91 GlnAlaLysLeuAlaAspLysLeuProIleProPheGlyGluSerValMetAspLeu 110
Db 498 TACTCCAAATTTCCGATAGCTGCGTCAAAATATCAGATGCCAATGAACCTGCGGATATG 557
QY 111 ValValIleGlyCysGlyProAlaGlyLeuSerLeuAlaAlaGluAlaLysLeuGly 130
Db 558 GTGGTTATTGGTGTGTCAGCTGCTAGCTTTGGCTGCAGAAATCGGCTAACTTGA 617
QY 131 LeuLysValGlyLeuIleGlyProAspLeuProPheThrAsnAsnTyrGlyValTrpGlu 150
Db 618 TTAAGAGTTGGCTTGTGCTGCTGATCTCTCTTTACGAAATACACGCGGTTGGGAA 677
QY 151 AspGluPheLysAspLeuGlyLeuGluArgCysIleGluHisAlaTrpLysAspThrIle 170
Db 678 GATGAATTCAGACATTTGGAGCTTTGGAGCTGTATCGACACAGCTTTGGCGTGATACCAT 737
QY 171 ValTyrLeuAspAsnAspAlaProValLeuIleGlyArgAlaTyrGlyArgValSerArg 190
Db 738 GTGTATATTGATGATGACATCTATATATATTGGTCGATCTTATGGAAAAGTCAGCCGG 797
QY 191 HisLeuLeuHisGluGluLeuLeuLysArgCysValGluSerGlyValSerTyrLeuAsp 210
Db 798 CAATTACTTCAACAGGAACCTGGTGCACAGGTGTTGGAGTCAGGTGCTCTTATCTGAAT 857
QY 211 SerLysValGluArgIleThrGluAlaGlyAspGlyHisSerLeuValValCysGluAsn 230

```

```

Db 858 GCGAAAGTGGAAAAATATTATGGAAGAGCCTGATGGACATAGGCTTGTGCTTGTGAACGT 917
QY 231 GluIlePheIleProCysArgLeuAlaThrValAlaSerGlyAlaAaSerGlyLysLeu 250
Db 918 GGTGTCACTATTCCTCGAGGCTTGTAACTGTCATCTGGAGCAGCTTCAGGAAACTT 977
QY 251 LeuGluTyrGluValGlyGlyProArgValCysValGlnThrAlaTyrGlyValGluVal 270
Db 978 CTGGAGTATGAAGTGGGTGGTCCAAAGGTTTGTGTACAAACAGCTTATGTTGTGGAGGTG 1037
QY 271 GluValGluAsnAsnProTyrAspProAsnLeuMetValPheMetAspTyrArgAspTyr 290
Db 1038 GAGTGGAAAAACAGCTCTTATGATCCCAATGTATGTTGTTCATGGACTACAGACTAC 1097
QY 291 MetGlnGlnLysLeuGlnCysSerGluGluGluTyrProThrPheLeuTyrValMetPro 310
Db 1098 ACTAACTGAGCGTTCATCTCTGGAGCAAGATATCCAACTTCTGTATGCAATGCGG 1157
QY 311 MetSerProThrArgLeuPhePheGluGluThrCysLeuAlaSerLysAspAlaMetPro 330
Db 1158 ATATCACCACACTAGGACTTCTTTGAGGAGACTTGTCTGGCTTCAGTAGATGCAATGCC 1217
QY 331 PheAspLeuLeuLysArgLysLeuMetSerArgLeuLysThrLeuGlyIleGlnValThr 350
Db 1218 TTTGACCTGCTCAAGAAAAAGCTTATGACAAAGATTACAACTATGGGTGTTGCTATCAC 1277
QY 351 LysValTyrGluGluTyrSerTyrIleProValGlyGlySerLeuProAsnThrGlu 370
Db 1278 AATATATATGAGAGGAGTGGTCTTATATACCTGTTGGTGGTCTTACCAANATACAGAG 1337
QY 371 GlnLysAsnLeuAlaPheGlyAlaAlaAaSerMetValHisProAlaThrGlyTyrSer 390
Db 1338 CAAAGAAACCTTGCACTTTGGTGTCTGCGAGCATGGTGCATCCAGCCACAGGTATTCA 1397
QY 391 ValValArgSerLeuSerGluAlaProLysTyrAlaSerValIleAlaLysIleLeuLys 410
Db 1398 GTCGTGAGATCACTGTCCAGAGCTCCAAAGTATGCTTCTGCAATGCAAACTGATCAAG 1457
QY 411 GlnAspAsnSerAlaTyrValValSerGlyGlnSerAlaValAlaAsnIleSerMetGln 430
Db 1458 AATGACCTGTCAAAAAATGCAATATTCGTCAGAGGAGTGTGGGGAATATCTCAATGCA 1517
QY 431 AlaTrpSerSerLeuTyrProLysGluArgLysArgGlnArgAlaPhePheLeuPheGly 450
Db 1518 GCCTGGAAATACTTCTTGGCCACAGAAAGAAAGCTCAGAGAGCATTTCTCTGTTGCGA 1577
QY 451 LeuCluLeuIleValGlnLeuAspIleGluAlaThrArgThrPheArgThrPhePhe 470
Db 1578 CTATCACTTATAGTCCAGCTTGTATGAGGGTATTCAGGACATTTCTCCGACCTTCTTC 1637
QY 471 ArgLeuProThrTrpMetTrpTyrPheLeuGlySerSerLeuSerSerPheAspLeu 490
Db 1638 CGAGTGCACAAATGATGTGGGAGGATTCCTCGTCTTATCTCTCTTACCTGATCTC 1697
QY 491 ValLeuPheSerMetTyrMetPheValLeuAlaProAsnSerMetArgMetSerLeuVal 510
Db 1698 ATATTGTTGCTTTTATATGTTCTTTATGCTCCGATGACTTGCAGAAATGGTCTTATA 1757
QY 511 ArgHisLeuLeuSerAspProSerGlyAlaValMetValArgAlaTyrLeu 527
Db 1758 AGGCATCTACTATCTGATCTTACAGGGGCGGACCATGATGAAGACGTACATA 1808

```

RESULT 13
 AF486650 1550 bp mRNA linear PLN 21-MAR-2002
 LOCUS Citrus x paradisi lycopene epsilon-cyclase mRNA, complete cds.
 DEFINITION AF486650
 ACCESSION AF486650
 VERSION AF486650.1 GI:19569600
 KEYWORDS Citrus x paradisi.
 SOURCE Citrus x paradisi
 ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eurosids II; Sapindales; Rutaceae; Citrus.
1 (bases 1 to 1550)
REFERENCE Costa,M.C., Otoni,W.C. and Moore,G.A.
AUTHORS Direct Submission
TITLE Submitted (22-FEB-2002) Horticultural Sciences, University of
JOURNAL Florida, 1301 Riefeld Hall, Gainesville, FL 32611, USA
FEATURES Location/Qualifiers
source 1..1550
/organism="Citrus x paradisi"
/cultivar="Duncan"
/db_xref="taxon:37656"
CDS 45..1358
/codon_start=1
/product="lycopenase epsilon-cyclase"
/protein_id="AAU92114.1"
/db_xref="GI:19569601"
/translation="MDMFLPIPSITGILVDVIGCGPAGLALAAESAKILGNVLGI
EDLLRCPNNYGWDEFDLGLGECIEHWRDVTVYIDDEPILIGRAYGRVSRHLH
BEELRRCEVSQYLSKVESIESTSGHRIVACEHDIMVPCRLATVASGAASGLKE
YEVGGPKSVQTAQVEYVENNVPDYRMFNDYRDCTQEVPFSFSDNPFTLIYMP
MSSTVFETCTACKDGRLPDILKKMLARLELRIGLOVLYTEEYSIPVGGSPLN
TBORNLTGAASAASHVPAQTGVSVRSISEAPNSAIAIILKHDSHSRGTLTHQSNE
ISKQWNTLPQERKRORAFRELGALILQDIDEGIRTFRTFFTRLPKMMWHGFLOGSS
LSSADILLFAFMFIANDRKLIKRLHLSVDPDTGATMVRYTLT"
BASE COUNT 469 a 270 c 333 g 478 t
ORIGIN

Alignment Scores:		
Pred. No.:	7,49e-144	Length: 1550
Score:	1786.00	Matches: 333
Percent Similarity:	87.67%	Conservative: 44
Best Local Similarity:	77.44%	Mismatches: 53
Query Match:	65.11%	Indels: 0
DB:	8	Gaps: 0

US-09-701-395A-23 (1-529) x AF486650 (1-1550)

QY	98	LeuProPheIleProPheGlyClusSerValMetAspLeuValValIleGlyCysGlyPro	117
Db	60	CTACGCCCAATCATTTGGTAATGGTATTTGGATTGGTGTTGGTGTGGCCCA	119
QY	118	AlaGlyLeuSerLeuAlaAlaAlaLysLeuGlyLeuLysValGlyLeuIleGly	137
Db	120	GCTGTCTTGTCTTGTGGTCGAGAATGGGATTAATAATGTTGCATTATGGC	179
QY	138	ProAspLeuProPheThrAsnAsnTyrcGlyValTrpGluaspGluPheLysAspLeuGly	157
Db	180	CGGATCTCCCTTTCACAACAACATATGGTGTGGGAAGATGAATTTAGAGATCTTGA	239
QY	158	LeuGluArgCysIleGluHisAlaTrpLysAspThrIleValTyrLeuAspAsnAla	177
Db	240	CTTGAAGGGTGTATCGAACATGCTCGAGAGACACACTTGTATATATTGATGAGATGA	299
QY	178	ProValLeuIleGlyArgAlaTyrGlyArgValSerArgHisLeuHisGluGluLeu	197
Db	300	CCCATCTTGATGTCGTCTTATGGACGAGTAGTCGACATTTGCTCATGAAGAATTA	359
QY	198	LeuLysArgCysValGluSerGlyValSerTyrLeuaspSerLysValGluArgIleThr	217
Db	360	TTAAGAAGGTGTGTCGAGTCAGGTGTTTCATATCTTAGCTCAAAGTGAAGACATTAG	419
QY	218	GluAlaGlyAspGlyHisSerLeuValValCysGluasnGluilePheIleProCysArg	237
Db	420	GAATCTACCAGTGGTCATCGCTGTGTAGCTGTGGAACATGATGATTTGTCCTTCAGG	479
QY	238	LeuAlaThrValIalaSerGlyAlaAlaSerGlyLysLeuLeuGluTyrGluValGlyGly	257
Db	480	CTTGCTACTGTGCTTCGAGCAGCATCAGGAAGCTATTGGAATATAGGTGGGGGT	539
QY	258	ProArgValCysValGlnThrIlaTyrGlyValGluValGluValGluasnAsnProTyr	277
Db	540	CCCCAAGTTTCGTCCAACACACTTAAGTGTGGAGGTGAGGTGAAAAATAATCCCATAT	599

Db	1390	ATCAGCTAACGCTTGGGATTAAAGTTGGACTCATTTGGTCCAGATCTCTCTTTTACTAACAA	14149
Qy	145	nTyrGlyValTrpGluAspGluPhe-----	153
Db	1450	TTACGGTGTTTGGGAAGATCAATTCAATGGTAAAAATCTCAATGATCCATTGTTTATGTC	15099
Qy	154	-----LysAspLeuGlyLeu	158
Db	1510	ACTATCTTCATGGTCTGTGTTTAAACATGTTGTTTATTGTTTACTGTCTGAAGATCTTTGGGCT	15699
Qy	158	uGluAtrGysIleuGluHisAlaTrpLysAspThrIleValTyrLeuAspAsnAspAlaPr	178
Db	1570	GCAAAATGTTATGACATGTTTGGAGAGAGACTATGTGTATCTGGATGATGACAAACC	16299
Qy	178	oValLeuIleGlyArgAlaTyrGlyArgValSerArgHisLeuLeuHisGluLeuLeu	198
Db	1630	TATTACCATTGGCCGCTTATGGAAGAGTTAGTCGACGTTTGTCTCCATGAGGAGCTTTT	16899
Qy	198	uLys-----	199
Db	1690	GAGGAGTAAAGTATGTTCCACTATGCATATCTTTGAGAGGCTTTTCTCTAGTTATTTT	17499
Qy	200	-----ArgCysValG	203
Db	1750	TGCATGACTTTGAAGCTTTTAAATACATAATTTGTTATTATATGCTTCTCAGGTGTGTCG	18099
Qy	203	luSerGlyValSerTyrLeuAspSerLysValGluArgIleThrGluAlaGlyAspGlyH	223
Db	1810	AGTCAGGTGTCGTACCTTAGCTCGAAAGTTGACAGCATAAACAGAAGCTTCTGATGCC	18699
Qy	223	isSerLeuValValCysGluAsnGluIlePheIleProCysArgLeu-----	238
Db	1870	TTAGACTTGTGTTGTGACGACAATAAGCTCATTCCTTGCAGGTA-ACAAAGTCAGCTAG	19289
Qy	239	-----AlaI	240

Db 1929 TACTATTCTGGTTTACCACCTACGAAAGGTTCTGACTTCTAGCAAATGCAGGCTTGCCA 1988

QY	240	nrValAlaSerGlyAlaAlaSerGlyLysLeuLeuGluTyrGluValGlyGlyProArgV	260
Db	1989	CTGTTGCTTCGGAGCAGCTTCGGGAAAGCTCTTGCATACGAACTTGCTGGACCTAGAG	2048
QY	260	alCysValGlnThrAlaTyrGlyValGluValGlu-----	271
Db	2049	TCYGTGTGCAACTGCATACGGCGTGGAGGTTAGAGTATAAATTGAATATGATATTCAC	2108
QY	271	-----	271
Db	2109	AGCAATAACTATATGAGGAAATCATGTTTCATTGAGACTGTTTTCTATTGAAATCTTC	2168
QY	272	--ValGluAsnAsnProTyrAspProAsnLeuMetValPheMetAspTyrArgAspTyrM	291
Db	2169	AGGTGGAAATATGCTCATATGATCCAGATCAAAATGGTTTTTCATGGATTACAGAGATTATA	2228
QY	291	etGlnGlnLysLeuGlnCysSerGluGluGluTyrProThrPheLeuTyrValMetProM	311
Db	2229	CTAACGAGAAAGTTCGGAGCTTAGAAGCTGAGTATCCCAACGTTCTCTGTACGCCATGCCTA	2288
QY	311	etSerProThrArgLeuPheGlu-----	319
Db	2289	TGACAAAGTCAAGACTCTTCTTCGAGGTTCTCTTTGGTGTCTCTCTCTCTAGTTAA	2348
QY	320	-----GluThrCysLeu	323
Db	2349	ATCGTTATTCGTATAAGTCTGTGCTCATCAGCTGGAATCCTTTTCAGGAGACATGTTTG	2408
QY	324	AlaSerLysAspAlaMetProPheAspLeuLeuLysArgLysLeuMetSerArgLeuLys	343
Db	2409	GCCTCAAAAGATGTCATGCGCCCTTGATTGCTTAAACCAAGCTCATGTTAAGATTAGAT	2468
QY	344	ThrLeuGlyIleGlnValThrLysValTyrGluGlu-----	355
Db	2469	ACACTCGGAATTCGAATCTCTAAAGACTTACGAAAGAGGTAAAGAAATGGTGGCTCTATATCA	2528

QY 356 -----GluTTPSer 358
 Db 2529 ACATAGAACTCTTATTGCTTTTAAATTCGTTTAAATATAAACTGTAGGAGTGGTCC 2588
 QY 359 TyrileProValGlyGlySerLeuProAsnThrGluGlnLysAsnLeuAlaPheGlyAla 378
 Db 2589 TATATCCAGTGTGGTGGTCTCTGCAACACCGAACAAGAAATCGCCTTGGTGTCT 2648
 QY 379 AlaAlaSerMetValHisProAla----- 386
 Db 2649 GCCGCTAGCATGTATACATCCGCAACAGAGTAGTAAAGCAAGAGCCTTGACCCGATGTA 2708
 QY 387 -----ThrG1 388
 Db 2709 CTTTTATTACATCTTGAATCTGTACCACTCTGACTTAGCTTTCCGCCAATTAACACAGG 2768
 QY 388 yTyrSerValAlaArgSerLeuSerGluAlaProLysTyrAlaSerValIleAlaLysI1 408
 Db 2769 CTATTCAGTGTGAGATCTTTCTGAGCTCCAAATATGATCAGTCATCGCAGAGAT 2828
 QY 408 eLeuLysGlnAspAsnSerAlaTyrValValSerGlyGlnSerSerAlaValAlaSnIleSe 428
 Db 2829 ACTAGAGAGAGAGACTACCAACAGATCAACAGT-----ANATATTC 2870
 QY 428 rMetGln----- 430
 Db 2871 AAGACAAGGTGAGCTCTATATAAACCAAGTTAACTATGCTCTTTTCACACAATGA 2930
 QY 431 -----AlaTrpSerSerLeuTrpProLy 438
 Db 2931 GAACTTTTAAACTGTGATTTTCTATTTTGGCAGCTTGGGATACTTTATGGCCACC 2990
 QY 438 sGluArgLysArgGlnArgAlaPhePheLeuPheGlyLeuGluLeuValGlnLeuAs 458
 Db 2991 AGAAGGAAAAAGACAGAGAGATCTTCTCTTGGCTTGCACATCATAGTTCATTCGA 3050
 QY 458 pileGluAlaThrArgThrPhePheArgThrPhePheArgLeuProThrTrp----- 475
 Db 3051 TACCAAGGAGATTAGAAGCTTCTCCGCTACTTCTCCGCTTCCAAATGGAACAG 3110
 QY 475 ----- 475
 Db 3111 ACATGATTGCTGATTATCAACAATGTTACTGTAAACAGAAATGATTCCTTTTCA 3170
 QY 476 -----MetTrpTrpGlyPheLeuGlySerSerLeuSerSerPheAspLeuValle 492
 Db 3171 CTTACTCAGATGTGGCAAGGTTCTAGGATCAACATTAACATCAGGAGATCTGTTCT 3230
 QY 492 uPheSerMetTyrMetPheValLeuAlaProAsnSerMetArgMetSerLeuValArgH1 512
 Db 3231 CTTTGTCTTATACATGCTGCTATTCACCAACAAATTTGAGAAAGGTCTCATCAATCA 3290
 QY 512 sLeuLeuSerAspProSerGlyAlaValMetValArgAlaTyrLeuGlu 528
 Db 3291 TCTCATCTCTGTATCAACCGGAGCAACCATGATAAAACCTATCTCAAA 3339
 RESULT 15
 AB024035
 LOCUS
 DEFINITION Arabidopsis thaliana 78423 bp DNA linear PLN 27-DEC-2000
 ACCESSION AB024035 BA000015
 VERSION AB024035.1 GI:4519194
 KEYWORDS Arabidopsis thaliana (strain:Columbia) DNA, clone_lib:Mitsui P1
 SOURCE clone:MHI7
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (sites)
 Sato, S., Nakamura, Y., Kaneko, T., Katoh, T., Asamizu, E., Kotani, H.
 and Tabata, S.

TITLE
 JOURNAL
 MEDLINE
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Structural analysis of Arabidopsis thaliana chromosome 5. x.
 Sequence features of the regions of 3,076,755 bp covered by sixty
 P1 and TAC clones
 DNA Res. 7 (1), 31-63 (2000)
 20181125
 2 (bases 1 to 78423)
 Nakamura, Y.
 Direct Submission
 Submitted (24-FEB-1999) Yasukazu Nakamura, Kazusa DNA Research
 Institute, Department of Plant Gene Research; 1532-3, Yana,
 Kisarazu, Chiba 292-0612, Japan (E-mail: ynakamuekazusa.or.jp,
 Tel: 81-438-52-3935, Fax: 81-438-52-3934)
 Address for correspondence: kaos@kazusa.or.jp
 For the latest information on annotation of this clone, please see
 http://www.kazusa.or.jp/kaos/cgi-bin/agd-graph.cgi?rc=MHI7
 Genes with similarity to proteins in the databases are described in
 'product' or 'note' qualifiers. Genes that have no significant
 protein similarity are described as 'unknown protein'.
 The software programs used to predict genes include: Grail
 (Informatics Group, Oak Ridge National Laboratory,
 http://compbio.ornl.gov/Grail-1.3/),
 GENSCAN (S.M. Hebsgaard, et al., http://CCR-081.mit.edu/GENSCAN.html),
 NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of
 Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and
 SplicePredictor (Volker Brendel, Stanford University,
 http://gremilni.zool.iastate.edu/cgi-bin/sp.cgi).
 Genes encoding tRNAs are predicted by tRNAscan-SE
 (Sean Eddy, Washington University School of Medicine, St. Louis,
 http://genome.wustl.edu/eddy/tRNAscan-SE/).
 This sequence may not be the entire insert of this clone. It may be
 shorter because we remove overlaps between neighboring submissions.
 The 5' clone is MP110 and the 3' clone is MUL3.
 Location/Qualifiers
 1..78423
 /organism="Arabidopsis thaliana"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /chromosome="5"
 /clone="MHI7"
 /clone_lib="Mitsui P1"
 complement(1..30)
 /note="CDS is reported in Acc# AB020747
 contains similarity to protein kinase
 gene_id:MP110.5"
 /number=1
 /evidence=not_experimental
 complement(join(4221..4349,4616..4761,4851..4944,
 5022..5307,5397..5821,6091..6208,6305..6432,6529..6638,
 6709..6982,7070..7141))
 /note="dbj|BAA91947.1
 gene_id:MHI7.1
 similar to unknown protein"
 /codon_start=1
 /evidence=not_experimental
 /protein_id="BA097020.1"
 /db_xref="GI:8777430"
 /translation="MAPRIPLCTGDLRLNLFKRVQSVSKSAGFPDALICGQFFPD
 SPEILDFDYVGRQVPIPTFTGDYGVAPKLTSTSKAENQFKMGDLGVCHN
 LFLWRSQGLFSLHGLSVAYLSGROSSSEQFKYQDDVDALRALEPFGGCTNWEFA
 GVTNRAAVSDIPVGLSDSCSDSTVSELMVKRYHAGSNGVIFYAREPLYNAESTK
 VTRFTGLAQVGNKKQKLEHALSTPTSTSPASLSAKPKPTTLWPHLODGAESKK
 RPNVSDSDQIWRIDVPRKQSGSGQKLCFKVCSGCPRGDCHEFHQNAAREOCR
 RGVLDDLLIKGCKEPCSYKHEFQDESIQKPRSENANRSKECFCLSSPSVESH
 LTVSGESFYCALPKGLVEDHILIIPIEHLPTNLVLSPEVSELSYQNGLRNCYS
 QGNDVFFELYSKRVSHANLQVVPVPSRRALLNIFSLAAKLGFLGKLVKPKNDST
 NCKLYQKEYNALGLFVYELPDGVLVSHTEENEFVPAQGREVLGLLIPDRWR
 NKLSQREEAALAEKQFQEDPCQ"
 join(7467..7643,7741..8080,8177..8230,8376..8479)
 /note="gene_id:MHI7.2
 pir||C/1422
 similar to unknown protein"
 /codon_start=1
 /evidence=not_experimental

```

/protein_id="BAA97021.1"
/db_xref="GI:8777431"
/translation="MADSVLKEVNGCRPEKIPKLDKACGSKSWKHLKLGNEVDEDEY
KROYCLFVEFKSEGFVDWEKDYMFHLRPLENSPPTSDIRTNADVIRDLTFATE
KNEAGSKLVEVHVSAFKANGELLCWLTFWATDMASAPTSQVYVVELWRGQKF
EIPFPRKDEMDDDVEKPPSPMEYDDYDKPVPVVRAPEDGVPVFDRTGALID
LRSGI"
CDS
  join(9471..9653,9736..10236)
  /note="gene_id:MHM17.3
  pir||C71422
  similar to unknown protein"
/codon_start=1
/evidence-not_experimental
/protein_id="BAA97022.1"
/db_xref="GI:8777432"
/translation="MADSLMEVNYRRPEKIPKLDCEGSESSKORVVKWGTAEDD
EYLROYLLPHYQKQFSGSIKWEQDYNFRSMDKSPNEISERKSNVILIREMLIT
AIDKNEAGTKLVFVEHVEANYQLRGLTCWLTFWATDMASPPSKYIQAHWRG
QNFHFIKRLPDKDEIEAVEVQPPSPMLYELDKPPIVFSRPGDEALPGVPVFNR
TGAGLDPDW"
CDS
  complement(join(12420..13157,13254..13394,13492..13542,
  13628..13759,13862..13927,14005..14747,14880..15036))
  /note="gene_id:MHM17.4
  unknown protein"
/evidence-not_experimental
/codon_start=1
/protein_id="BAA97023.1"
/db_xref="GI:8777433"
/translation="MESSVAPPHRTHLPNRYRSLYHFCSSFDREPOISLATPAY
LQELVLRTEIAESPGETCEPPENLSITESKLVNGSDSGERKETISQKSLMIGDI
CGIDIQDASVSRHTDFDFELMNETQDSVESCYNLFEALDNDYDIDVQNVLEK
PNATQVQVSEKAEKVPKSENEVISSGVLEACNGTVQREMELEKPKVDNSPV
LVDVSRIVGGDDVEGEISGDNDMLVEDDTERHEEYQVSGDTGNSHLTSHKS
FGVEVMNDVQAQKIDQTSNEAKMDPGTSIKRSAPSDAKAKRAKAKIRAQORI
ALGVKLLKVPAPKPKIKYCHRYLKGCHGEGDKCFSHDTIPFKSCPCCFATQS
CMKGDCCPDHLSKXPCNFTKGYRGDSCLFSGHGTQPSASDTPSANVTVSSTK
ITAAESPQKQSVRDAIAKLPAIQARVSSVAFKFPSSHSNRNSDASSKINE
HYTPQVPLRPKPSVAPKMSFLSDKTSQEDTVKASSAKSPNTDSDQTLKQSQG
SFLPLGPPKGISFLSPASEEQTLNRPQKPPASSKMLKTPSSHIOSSLISAMKLA
FESAKVERGNNDPTEAVNKSNTVDVATRNNGNISSKILEFLSSFSHGKN"
  join(15350..15482,15570..15615,16443..16671)
  /note="gene_id:MHM17.5"
/codon_start=1
/evidence-not_experimental
/product="30S ribosomal protein S16"
/protein_id="BAA97024.1"
/db_xref="GI:8777434"
/translation="MVVRIILSRFGCKNRPFFRYMAADSRSPRDKHLEVLGFENLPL
GDGGRMGLKFDRIKYLWSVGQPSDPVORLLFRSGLLPPPPMVAWGRKGARTRP
VDPMTGRYVDAENKTVNANDNQPEEDTEAKS"
  join(17089..17134,17217..17260,17389..17484,17694..17786,
  17919..18006,18095..18156,18245..18429,18562..18739,
  18827..18870,19017..19218,19684..19770)
  /note="gene_id:MHM17.6"
/codon_start=1
/evidence-not_experimental
/product="nucleosome assembly protein"
/protein_id="BAA97025.1"
/db_xref="GI:8777435"
/translation="MSNDKDSFNVSDLTSALKDDEDRAGLVNALKNLQNLAGHSDVL
ENITKIRRVLRVLRIOGHDEIEFKFREAALEAKYOKLYQPLNKRKYEVNAGT
EVEGAPAKMDQGDKEAERGVSPFWLTALKNDVISEITEFERDEGALYLKLDIKW
CKIEPKGLFEFFQDNPYFKNTLTKAYHIDEDEPLELKEAIGTEIDWYPKCLTQ
KILKKPKKAKNAKPIKTDECSFEFFNFNPQVDDDEDIDEERAEELQNLMEODY
DIGSTIREKLIIPHVSFWTGEALGEEFEEDDDEDDEDDEDEDEDEDEDEDEDE
DREBEVSKTKKFEIILQKGRPVQVDDQGERPPECKQQ"
  complement(join(23174..23239,23666..24049,24174..24863,
  25283..25388,25491..25645))
  /note="gb|AAC78547.1
  gene_id:MHM17.7
  similar to unknown protein"
/codon_start=1
/evidence-not_experimental

```

```

/protein_id="BAA97026.1"
/db_xref="GI:8777436"
/translation="MMHLILCSLYLISMDGYVNEASEPSSSSSSSSSLARSILFHEYRO
SVIPLQNGHVPSPMAFMNMLPVEIRPOESQRLAFNDLQRLFYQNKIEASLEWPEDEP
NRKSPANSOYLRPHYPSPSSSSSSLSPPNISEYSSLLPLPKPSTTTEAVNVVPLPP
LAPINMHPHQEPLFRNRQREEMTQAILAVLTGSPSPSTSPQKRGATAFKR
YYSMISDRAPLPMVRKQSMTRAMSFNRLNINQREFTRENATTHGEGSGGG
GRYTSGPSATQLOHMISEKREKLNSLPCTCKDKASVLSIAREQSSSL
QGEISKLERNREVEAKIAGEREINDLRPERFNRVRIHPIESTSRERTLDLRVLRL
GDITRVDLMLRLEFLQKINNVSLSVEARTLARAEGDTSIVLISLRLEGEWDE
SAFQEAERVVVADLAH"
CDS
  complement(join(29415..29718,29813..30081,30183..30443,
  30529..30656,32107..32716))
  /note="gene_id:MHM17.8"
/codon_start=1
/evidence-not_experimental
/product="cytokinin oxidase"
/protein_id="BAA97027.1"
/db_xref="GI:8777437"
/translation="MASYNLRSQVRLIAITVIIITLSTPTTNTSPQWNILSHNEF
AGHSHRGOASAKDGVVNMRSVMNRDGIKVSRTCLYVDVAAMWIEVLAKTLLEGL
KCTSSSSSVESAATDFGHVTKIPPSAVLIPSSVEDITDLIKLSDQSPPLAARG
HGHSHRGOASAKDGVVNMRSVMNRDGIKVSRTCLYVDVAAMWIEVLAKTLLEGL
TPVSWTDYLYLTVGGLSNGSIGQTFERYQIYNLEMDVITGKEATATSKDNMSD
LFFAVLGGGQFGIITRARIKLEAPRAKWLRLYIDFSETRDQERIVISKTGDVDF
LEGSIMVDHGPPDNMRSTYPPSDHLRIASAKVHRVLYCLEVVKYIDETSYTVNEE
MEELSDSLNHRVGFMYEKDVTYMDFLNRVTCGLNLSKQMDVPPHVLNLFVPRQI
SKPDDGVFKGIILRNNTISGPVLVPMNRNKNDRMSAAIPEEDVYVAVGLRSAGFD
NWEAFDOENMEILAFCEADANNVIOQLPYHSSQEGVWRHFGPRNIVFVKYKIDPKM
ILSPQNIQKINSS"
  complement(39239..39315)
  /product="tRNA-Val"
  /note="codon recognized: GUG; gene_id:MHM17.9"
  /evidence-not_experimental
  complement(join(41045..41748,41932..42184))
  /note="gene_id:MHM17.10
  pir||T04268
  similar to unknown protein"
/codon_start=1
/evidence-not_experimental
/protein_id="BAA97028.2"
Alignment Scores:
Pred. No.: 1,54e-128 Length: 78423
Score: 1633.50 Matches: 376
Percent Similarity: 53.36% Conservative: 61
Best Local Similarity: 45.91% Mismatches: 80
Query Match: 59.55% Indels: 303
DB: 8 Gaps: 15

US-09-701-395A-23 (1-529) x AB024035 (1-78423)
QY 1 MetGluLeuLeuGlyValArgAsnLeu-----IleSerSerCysProValTyr 16
Db 61869 ATGGAGTGTGTGGGGCTAGGAATTTCCGACCAATGGCGGTTCACATTCGTCATGG 61928

QY 17 ThrPheGlyThrArgAsnLeuSerSerSerLysLeuAlaTyrAsnIleHisArgTyrGly 36
Db 61929 AGTTGTCCAGGAAGAAATTTCCAGTGTGTTAAAGATACAGCATAGGAATATTCGTTCCGT 61988

QY 37 SerSerCysArgValAspPheClnValArgAlaAspGlyGly-----SerGlySerArg 54
Db 61989 --TTGTGT-----AGTGTACAGCTAGCGCGGGAAGTTCCGCTAGTAGAG 62033

QY 55 SerSerValAlaTyrLysGluGlyPheValAspGluGluAspPheIleLysAlaGlyGly 74
Db 62034 AGTTGTGAGCGGTGAGAGAGATTTCCCTGACGAAGAAGATTTTGTGAAACCTGGTGT 62093

QY 75 SerGluLeuLeuPheValClnMetGlnGlnThrLysSerMetGluLysGlnAlaLysLeu 94
Db 62094 TCTGAGATTCATTGTTCAATGCGAGACAAAGATATGATGATGACAGCTTAAGCTT 62153

QY 95 AlaAspLys----- 97

```


GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 21, 2003, 22:11:13 ; Search time 189 Seconds
(without alignments)
277.579 Million cell updates/sec

Title: US-09-701-395A-23

Perfect score: 2743

Sequence: 1 MELLGVRNLISCPVWTRGT.....VRHLISDPGAVMVRAYLER 529

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 375593 seqs, 99172665 residues

Total number of hits satisfying chosen parameters: 375593

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2743	100.0	529	10	US-09-323-998D-23
2	2743	100.0	529	10	US-09-323-998D-50
3	2737	99.8	529	10	US-09-323-998D-47
4	2728	99.5	529	10	US-09-323-998D-51
5	1959.5	71.4	524	10	US-09-323-998D-2
6	1959.5	71.4	524	10	US-09-323-998D-21
7	1959.5	71.4	524	10	US-09-323-998D-49
8	1934	70.5	533	10	US-09-323-998D-52
9	1934	70.5	533	10	US-09-323-998D-54
10	1909.5	69.6	526	10	US-09-323-998D-53
11	1882	68.6	533	10	US-09-323-998D-26
12	1518.5	55.4	378	10	US-09-323-998D-25
13	1518.5	55.4	378	10	US-09-323-998D-25
14	1513	55.2	374	10	US-09-323-998D-27
15	775.5	28.3	502	10	US-09-323-998D-56
16	761	27.7	498	10	US-09-323-998D-57
17	755	27.5	511	10	US-09-323-998D-60
18	749.5	27.3	500	10	US-09-323-998D-58
19	748	27.3	501	10	US-09-323-998D-55

20	737	26.9	500	10	US-09-323-998D-59	Sequence 59, Appl
21	701.5	25.6	456	10	US-09-323-998D-20	Sequence 20, Appl
22	701.5	25.6	503	10	US-09-323-998D-61	Sequence 61, Appl
23	129.5	4.7	382	9	US-09-941-947A-30	Sequence 30, Appl
24	115.5	4.2	424	9	US-09-738-626-4029	Sequence 4029, Ap
25	113	4.1	382	9	US-09-920-923-5	Sequence 5, Appl1
26	113	4.1	382	10	US-09-547-267-7	Sequence 7, Appl1
27	103	3.8	433	10	US-09-815-242-5893	Sequence 5893, Ap
28	103	3.8	574	9	US-10-025-380-1107	Sequence 1107, Ap
29	103	3.8	574	10	US-09-922-217-1107	Sequence 1107, Ap
30	101	3.7	435	10	US-09-815-242-10696	Sequence 10696, A
31	99.5	3.6	567	10	US-09-815-242-13563	Sequence 13563, A
32	98	3.6	392	10	US-09-815-242-11795	Sequence 11795, A
33	98	3.6	435	10	US-09-815-242-12859	Sequence 12859, A
34	96.5	3.5	1146	10	US-09-824-734-2	Sequence 2, Appl1
35	95	3.5	419	10	US-09-815-242-13798	Sequence 13798, A
36	93.5	3.4	722	9	US-09-906-419-19	Sequence 19, Appl
37	93	3.4	469	9	US-09-738-626-4264	Sequence 4264, Ap
38	92	3.4	392	10	US-09-924-256A-24	Sequence 24, Appl
39	92	3.4	548	9	US-09-738-626-6486	Sequence 6486, Ap
40	92	3.4	751	10	US-09-815-242-5832	Sequence 5832, Ap
41	92	3.4	825	10	US-09-815-242-12963	Sequence 12963, A
42	91.5	3.3	799	10	US-09-952-677-6	Sequence 6, Appl1
43	89.5	3.3	1140	9	US-10-108-605-293	Sequence 293, App
44	89	3.2	400	10	US-09-815-242-10380	Sequence 10380, A
45	89	3.2	451	10	US-09-815-242-13127	Sequence 13127, A

ALIGNMENTS

RESULT 1

US-09-323-998D-23

Sequence 23, Application US/093233998D

Patent No. US20020102631A1

GENERAL INFORMATION:

APPLICANT: CUNNINGHAM JR., FRANCIS X.

APPLICANT: SUN, ZAIREN

TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND

TITLE OF INVENTION: METHODS OF USE THEREOF

FILE REFERENCE: 108172-09019

CURRENT APPLICATION NUMBER: US/09/323,998D

CURRENT FILING DATE: 1999-06-02

PRIOR APPLICATION NUMBER: 09/088,724

PRIOR FILING DATE: 1998-06-02

PRIOR APPLICATION NUMBER: 09/088,725

PRIOR FILING DATE: 1998-06-02

PRIOR APPLICATION NUMBER: 08/937,155

PRIOR FILING DATE: 1997-09-25

PRIOR APPLICATION NUMBER: 08/624,125

PRIOR FILING DATE: 1996-03-29

NUMBER OF SEQ ID NOS: 61

SOFTWARE: Patent Ver. 2.1

SEQ ID NO 23

LENGTH: 529

TYPE: PRT

ORGANISM: Adonis palaeatina

US-09-323-998D-23

Query Match 100.0%; Score 2743; DB 10; Length 529;

Best Local Similarity 100.0%; Pred. No. 1.1e-247;

Matches 529; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MELLGVRNLISCPVWTRGTGRNLSSKLAYNHRYGSSCRVDFQVRADGGSSRSVAYK 60

|||||

Db 1 MELLGVRNLISCPVWTRGTGRNLSSKLAYNHRYGSSCRVDFQVRADGGSSRSVAYK 60

|||||

QY 61 EGVFDEDFTKAGSELLFVQMOTKSMKQAKLADLPPIPGESVMDLVVTCGPGAGL 120

|||||

Db 61 EGVFDEDFTKAGSELLFVQMOTKSMKQAKLADLPPIPGESVMDLVVTCGPGAGL 120

|||||

QY 121 SLAAEAAKLGKVLGIPDLPFTNNYQWDEDFKDLGLERCIEHAWKDTTVYLDNDAPVL 180

|||||

Db 121 SLAAEAKLGLKVLGPDLPFTNNYGVWEDEFKDLGLERCIEHAWKDTIVYLDNDAPVL 180
QY 181 IGRAYGRVSRHLLHEELLKRCVSGVSYLDSKVERITEAGDGHSLVVCENEIIFPCRLAT 240
Db 181 IGRAYGRVSRHLLHEELLKRCVSGVSYLDSKVERITEAGDGHSLVVCENEIIFPCRLAT 240
QY 241 VASGAASGKLLLEYEGGPRVCVQTAYGVEVEVENNPDNLMVFMVDYRDYMOOKLOCSEE 300
Db 241 VASGAASGKLLLEYEGGPRVCVQTAYGVEVEVENNPDNLMVFMVDYRDYMOOKLOCSEE 300
QY 301 EYPTFLYVMPSPTRLFFETCLASKDAMPFDLLKRLKLSRLKTLGIGQVTKVYEEWSYI 360
Db 301 EYPTFLYVMPSPTRLFFETCLASKDAMPFDLLKRLKLSRLKTLGIGQVTKVYEEWSYI 360
QY 361 PVGGSPLNTEQKNLAFGAASVHPATGYSVVRSLSEAPKYASVIAKILKQNSAYVYSG 420
Db 361 PVGGSPLNTEQKNLAFGAASVHPATGYSVVRSLSEAPKYASVIAKILKQNSAYVYSG 420
QY 421 QSSAVNISQAWSSLPWKERKQRAFFLGLGLIYVOLDIEATRTFFRFLPTWMMWGF 480
Db 421 QSSAVNISQAWSSLPWKERKQRAFFLGLGLIYVOLDIEATRTFFRFLPTWMMWGF 480
QY 481 LGSSLSFSDLVLFMYMFVLAPNSMRMSLVRLHLLSDPSGAVMVRAYLER 529
Db 481 LGSSLSFSDLVLFMYMFVLAPNSMRMSLVRLHLLSDPSGAVMVRAYLER 529

RESULT 2

US-09-323-998D-50

; Sequence 50, Application US/09323998D

; Patent No. US20020102631A1

; GENERAL INFORMATION:

; APPLICANT: CUNNINGHAM JR., FRANCIS X.

; APPLICANT: SUN, ZAIREN

; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND
; FILE REFERENCE: 108172-09019

; CURRENT APPLICATION NUMBER: US/09/323,998D

; CURRENT FILING DATE: 1999-06-02

; PRIOR APPLICATION NUMBER: 09/088,724

; PRIOR FILING DATE: 1998-06-02

; PRIOR FILING DATE: 1998-06-02

; PRIOR FILING DATE: 1997-09-25

; PRIOR FILING DATE: 1996-03-29

; NUMBER OF SEQ ID NOS: 61

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 50

; LENGTH: 529

; TYPE: PRT

; ORGANISM: Adonis palaeastina

US-09-323-998D-50

Query Match 100.0%; Score 2743; DB 10; Length 529;
Best Local Similarity 100.0%; Pred. No. 1.1e-247;
Matches 529; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MELLGVRNLISPCVWTFGTNRNLSSSLAYNIHRYGSSCRVDFQVRADGGSGSRSSVAYK 60
Db 1 MELLGVRNLISPCVWTFGTNRNLSSSLAYNIHRYGSSCRVDFQVRADGGSGSRSSVAYK 60
QY 61 EGFVDEEDFTKAGSSELLFVQMOTKSMKQAKLADKLPPIPFGESVMDLVVIGCGPAGL 120
Db 61 EGFVDEEDFTKAGSSELLFVQMOTKSMKQAKLADKLPPIPFGESVMDLVVIGCGPAGL 120
QY 121 SLAAEAKLGLKVLGPDLPFTNNYGVWEDEFKDLGLERCIEHAWKDTIVYLDNDAPVL 180
Db 121 SLAAEAKLGLKVLGPDLPFTNNYGVWEDEFKDLGLERCIEHAWKDTIVYLDNDAPVL 180
QY 181 IGRAYGRVSRHLLHEELLKRCVSGVSYLDSKVERITEAGDGHSLVVCENEIIFPCRLAT 240
Db 181 IGRAYGRVSRHLLHEELLKRCVSGVSYLDSKVERITEAGDGHSLVVCENEIIFPCRLAT 240

Db 181 IGRAYGRVSRHLLHEELLKRCVSGVSYLDSKVERITEAGDGHSLVVCENEIIFPCRLAT 240
QY 241 VASGAASGKLLLEYEGGPRVCVQTAYGVEVEVENNPDNLMVFMVDYRDYMOOKLOCSEE 300
Db 241 VASGAASGKLLLEYEGGPRVCVQTAYGVEVEVENNPDNLMVFMVDYRDYMOOKLOCSEE 300
QY 301 EYPTFLYVMPSPTRLFFETCLASKDAMPFDLLKRLKLSRLKTLGIGQVTKVYEEWSYI 360
Db 301 EYPTFLYVMPSPTRLFFETCLASKDAMPFDLLKRLKLSRLKTLGIGQVTKVYEEWSYI 360
QY 361 PVGGSPLNTEQKNLAFGAASVHPATGYSVVRSLSEAPKYASVIAKILKQNSAYVYSG 420
Db 361 PVGGSPLNTEQKNLAFGAASVHPATGYSVVRSLSEAPKYASVIAKILKQNSAYVYSG 420
QY 421 QSSAVNISQAWSSLPWKERKQRAFFLGLGLIYVOLDIEATRTFFRFLPTWMMWGF 480
Db 421 QSSAVNISQAWSSLPWKERKQRAFFLGLGLIYVOLDIEATRTFFRFLPTWMMWGF 480
QY 481 LGSSLSFSDLVLFMYMFVLAPNSMRMSLVRLHLLSDPSGAVMVRAYLER 529
Db 481 LGSSLSFSDLVLFMYMFVLAPNSMRMSLVRLHLLSDPSGAVMVRAYLER 529

RESULT 3

US-09-323-998D-47

; Sequence 47, Application US/09323998D

; Patent No. US20020102631A1

; GENERAL INFORMATION:

; APPLICANT: CUNNINGHAM JR., FRANCIS X.

; APPLICANT: SUN, ZAIREN

; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND
; FILE REFERENCE: 108172-09019

; CURRENT APPLICATION NUMBER: US/09/323,998D

; CURRENT FILING DATE: 1999-06-02

; PRIOR APPLICATION NUMBER: 09/088,724

; PRIOR FILING DATE: 1998-06-02

; PRIOR FILING DATE: 1998-06-02

; PRIOR FILING DATE: 1997-09-25

; PRIOR FILING DATE: 1996-03-29

; NUMBER OF SEQ ID NOS: 61

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 47

; LENGTH: 529

; TYPE: PRT

; ORGANISM: Adonis palaeastina

US-09-323-998D-47

Query Match 99.8%; Score 2737; DB 10; Length 529;
Best Local Similarity 99.6%; Pred. No. 4.1e-247;
Matches 527; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MELLGVRNLISPCVWTFGTNRNLSSSLAYNIHRYGSSCRVDFQVRADGGSGSRSSVAYK 60
Db 1 MELLGVRNLISPCVWTFGTNRNLSSSLAYNIHRYGSSCRVDFQVRADGGSGSRSSVAYK 60
QY 61 EGFVDEEDFTKAGSSELLFVQMOTKSMKQAKLADKLPPIPFGESVMDLVVIGCGPAGL 120
Db 61 EGFVDEEDFTKAGSSELLFVQMOTKSMKQAKLADKLPPIPFGESVMDLVVIGCGPAGL 120
QY 121 SLAAEAKLGLKVLGPDLPFTNNYGVWEDEFKDLGLERCIEHAWKDTIVYLDNDAPVL 180
Db 121 SLAAEAKLGLKVLGPDLPFTNNYGVWEDEFKDLGLERCIEHAWKDTIVYLDNDAPVL 180
QY 181 IGRAYGRVSRHLLHEELLKRCVSGVSYLDSKVERITEAGDGHSLVVCENEIIFPCRLAT 240
Db 181 IGRAYGRVSRHLLHEELLKRCVSGVSYLDSKVERITEAGDGHSLVVCENEIIFPCRLAT 240
QY 241 VASGAASGKLLLEYEGGPRVCVQTAYGVEVEVENNPDNLMVFMVDYRDYMOOKLOCSEE 300
Db 241 VASGAASGKLLLEYEGGPRVCVQTAYGVEVEVENNPDNLMVFMVDYRDYMOOKLOCSEE 300

Db 241 VASGAASGLLEYEGVGRVCQTAYGVEVEENNPDNLMVMDYRDYMOQKLOCSEE 300
QY 301 EYPTFLYVMPMSPTLFFETCLASDKAMPDGLLKKRLSKLTGLGIVTKYVEEWSYI 360
Db 301 EYPTFLYVMPMSPTLFFETCLASDKAMPDGLLKKRLSKLTGLGIVTKYVEEWSYI 360
QY 361 PVGGSPLNTEQKNLAFGAASVHPATGYSVVRSLSEAPKYASIAKILKQDNSAYVWSG 420
Db 361 PVGGSPLNTEQKNLAFGAASVHPATGYSVVRSLSEAPKYASIAKILKQDNSAYVWSG 420
QY 421 QSSAVNISMQAWSLWPKERKQRAFFLFGLELIVOLDIEATRTFFRFLPTMMWGF 480
Db 421 QSSAVNISMQAWSLWPKERKQRAFFLFGLELIVOLDIEATRTFFRFLPTMMWGF 480
QY 481 LGSSLSFDLVLFMTMFLAPNSMRSLVRHLLSDPSGAVMVVAYLER 529
Db 481 LGSSLSFDLVLFMTMFLAPNSMRSLVRHLLSDPSGAVMVVAYLER 529

RESULT 4
US-09-323-998D-51
; Sequence 51, Application US/09323998D
; Patent No. US20020102631A1
; GENERAL INFORMATION:
; APPLICANT: CUNNINGHAM JR., FRANCIS X.
; APPLICANT: SUN, ZAIREN
; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND
; FILE REFERENCE: 108172-09019
; CURRENT APPLICATION NUMBER: US/09/323,998D
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 09/088,724
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 09/088,725
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 08/937,155
; PRIOR FILING DATE: 1997-09-25
; PRIOR APPLICATION NUMBER: 08/624,125
; PRIOR FILING DATE: 1996-03-29
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 51
; LENGTH: 529
; TYPE: PRT
; ORGANISM: Adonis palaestina
US-09-323-998D-51

Query Match 99.5%; Score 2728; DB 10; Length 529;
Best Local Similarity 99.1%; Pred. No. 2.9e-246;
Matches 524; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
QY 1 MELLGVRLNLISSCPVMTFGRNLSSSKLAYNTHRYGSSCRVDQVVRADGSGSRSSVAYK 60
Db 1 MELLGVRLNLISSCPVMTFGRNLSSSKLAYNTHRYGSSCRVDQVVRADGSGSRSSVAYK 60
QY 61 EGFVDEEDFIKAGGSELLFVQMQTKSMEKQAKLADKLPIPPGESVMDLVVIGCGPAGL 120
Db 61 EGFVDEEDFIKAGGSELLFVQMQTKSMEKQAKLADKLPIPPGESVMDLVVIGCGPAGL 120
QY 121 SLAAEAAKGLKVLGIGLDPPTNNYGVWEDEFKDLGLERCIEHAWKDTIYLDNDAPVL 180
Db 121 SLAAEAAKGLKVLGIGLDPPTNNYGVWEDEFKDLGLERCIEHAWKDTIYLDNDAPVL 180
QY 181 IGRAYGRVSRHLLHEELLKRCVSGSYLDSKVERITEAGDGHSLVVCENEIFIPCLAT 240
Db 181 IGRAYGRVSRHLLHEELLKRCVSGSYLDSKVERITEAGDGHSLVVCENEIFIPCLAT 240
QY 241 VASGAASGLLEYEGVGRVCQTAYGVEVEENNPDNLMVMDYRDYMOQKLOCSEE 300
Db 241 VASGAASGLLEYEGVGRVCQTAYGVEVEENNPDNLMVMDYRDYMOQKLOCSEE 300
QY 301 EYPTFLYVMPMSPTLFFETCLASDKAMPDGLLKKRLSKLTGLGIVTKYVEEWSYI 360

Db 301 EYPTFLYVMPMSPTLFFETCLASDKAMPDGLLKKRLSKLTGLGIVTKYVEEWSYI 360
QY 361 PVGGSPLNTEQKNLAFGAASVHPATGYSVVRSLSEAPKYASIAKILKQDNSAYVWSG 420
Db 361 PVGGSPLNTEQKNLAFGAASVHPATGYSVVRSLSEAPKYASIAKILKQDNSAYVWSG 420
QY 421 QSSAVNISMQAWSLWPKERKQRAFFLFGLELIVOLDIEATRTFFRFLPTMMWGF 480
Db 421 QSSAVNISMQAWSLWPKERKQRAFFLFGLELIVOLDIEATRTFFRFLPTMMWGF 480
QY 481 LGSSLSFDLVLFMTMFLAPNSMRSLVRHLLSDPSGAVMVVAYLER 529
Db 481 LGSSLSFDLVLFMTMFLAPNSMRSLVRHLLSDPSGAVMVVAYLER 529

RESULT 5
US-09-323-998D-2
; Sequence 2, Application US/09323998D
; Patent No. US20020102631A1
; GENERAL INFORMATION:
; APPLICANT: CUNNINGHAM JR., FRANCIS X.
; APPLICANT: SUN, ZAIREN
; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND
; FILE REFERENCE: 108172-09019
; CURRENT APPLICATION NUMBER: US/09/323,998D
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 09/088,724
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 09/088,725
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 08/937,155
; PRIOR FILING DATE: 1997-09-25
; PRIOR APPLICATION NUMBER: 08/624,125
; PRIOR FILING DATE: 1996-03-29
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-323-998D-2

Query Match 71.4%; Score 1959.5; DB 10; Length 524;
Best Local Similarity 70.4%; Pred. No. 1.9e-174;
Matches 376; Conservative 60; Mismatches 81; Indels 17; Gaps 5;
QY 1 MELLGVRLNLISSCPVMTFGRNLSSSKLAYNTHRYGSSCRVDQVVRADGSGSRSSVAYK 54
Db 1 MECVGARNAFAAMAVSTFPSSCRKRKPPVVKRYSYRNIRFG-LC---SVRASGGSGSGSE 55
QY 55 SVYAYKEGFVDEEDFIKAGGSELLFVQMQTKSMEKQAKLADKLPIPPGESVMDLVVIG 114
Db 56 SCVAVRDEDFEEDFVQMQTKSMEKQAKLADKLPIPPGESVMDLVVIG 115
QY 115 CGPAGLSLAAEAAKGLKVLGIGLDPPTNNYGVWEDEFKDLGLERCIEHAWKDTIYLD 174
Db 116 CGPAGLSLAAEAAKGLKVLGIGLDPPTNNYGVWEDEFKDLGLERCIEHAWKDTIYLD 175
QY 175 NDAPVLIGRAYGRVSRHLLHEELLKRCVSGSYLDSKVERITEAGDGHSLVVCENEIF 234
Db 176 DDKPITIGRAYGRVSRHLLHEELLKRCVSGSYLDSKVERITEAGDGHSLVVCENEIF 235
QY 235 PCRLATVASGAASGLLEYEGVGRVCQTAYGVEVEENNPDNLMVMDYRDYMOQK 294
Db 236 PCRLATVASGAASGLLEYEGVGRVCQTAYGVEVEENNPDNLMVMDYRDYMOQK 295
QY 295 LOCSESEYPTFLYVMPMSPTLFFETCLASDKAMPDGLLKKRLSKLTGLGIVTKYVE 354
Db 296 VRSLEAEYPTFLYVMPMSPTLFFETCLASDKAMPDGLLKKRLSKLTGLGIVTKYVE 355
QY 355 EWSYIPVGGSLPNTQKNLAFGAASVHPATGYSVVRSLSEAPKYASIAKILKQDNS 414

Db 356 EWSYIPVGGSLPNTQKNLAFGAAASWHPATGYSVRSLSSEAPKYASVIAEILREET 415
QY 415 AYVVSQSSAVNISMQAWSLWPKRKQRAFFLFGLELIVOLDIEATRTFFTRFLPT 474
Db 416 KOINS-----NISQAWDTLWPPERKQRAFFLFGLELIVQFDEGIRSFTRFLPK 469
QY 475 WNWGFLGSSLSFDLVLFSTMTFVLAPNSMRSLVRHLLSDPSGAVMVRAYLE 528
Db 470 WNWGFLGSLTSGDLVLFALYMFVISPNNLRKGLNHLISDPTGATMIKTYLK 523

RESULT 6
US-09-323-998D-21
; Sequence 21, Application US/09323998D
; Patent No. US20020102631A1
; GENERAL INFORMATION:
; APPLICANT: CUNNINGHAM JR., FRANCIS X.
; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND
; FILE REFERENCE: 108172-09019
; CURRENT APPLICATION NUMBER: US/09/323,998D
; PRIOR FILING DATE: 1999-06-02
; PRIOR FILING DATE: 1998-06-02
; PRIOR FILING DATE: 1998-06-02
; PRIOR FILING DATE: 1997-09-25
; PRIOR FILING DATE: 1996-03-29
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-323-998D-21

Query Match 71.4%; Score 1959.5; DB 10; Length 524;
Best Local Similarity 70.4%; Pred. No. 1.9e-174;
Matches 376; Conservative 60; Mismatches 81; Indels 17; Gaps 5;

QY 1 MELLGVRNL-----ISSCPVMTFTGRNLSKSLAYNIHRYGSSCRVDFQVRADGG--SGSR 54
Db 1 MECVGARNFAMAVSTPFSKCRKFPVKRYISYRNIRFG--LC-----SVRASGGSSGE 55
QY 55 SSVAYKEGFVDEEDFIKAGSELFFVQMOKTSMKQAKLADKLPIPPFGESVMDLVVIG 114
Db 56 SCVAVREDFADEEDFVKAGSEILFFVQMOKMDQSKLVKLPPIISIGDGLDHWVIG 115
QY 115 CGPAGLSLAAEAAKGLKVLGDLPTFTNNYGVWDEDFDKGLERCIEHAWKDTIVYLD 174
Db 116 CGPAGLALAAESAKGLKVLGDLPTFTNNYGVWDEDFDKGLERCIEHAWKDTIVYLD 175
QY 175 NDAPVLITGRAYGRVSRHLLHEELLKRCVSGSVYLSKVERITEAGDGHSLVVCENEIFI 234
Db 176 DDKPITIGRAYGRVSRHLLHEELLKRCVSGSVYLSKVDISITEASDGLRLVACDNNVI 235
QY 235 PCRLATVASGAAGSKLLEYEVGGPRVCVQYATGYVEVEVNNPYDPNLMVPMYRDYMOOK 294
Db 236 PCRLATVASGAAGSKLLEYEVGGPRVCVQYATGYVEVEVNNPYDPNLMVPMYRDYMOOK 295
QY 295 LOCSEEEYPTFLVMPMSPTFLFEETCTCLASKDAMPFDLLKRLKMSRLKTLGTVKTYE 354
Db 296 VRSLEAEYPTFLVMPMSPTFLFEETCTCLASKDAMPFDLLKRLKMSRLKTLGTVKTYE 355
QY 355 EWSYIPVGGSLPNTQKNLAFGAAASWHPATGYSVRSLSSEAPKYASVIAEILREET 414
Db 356 EWSYIPVGGSLPNTQKNLAFGAAASWHPATGYSVRSLSSEAPKYASVIAEILREET 415
QY 415 AYVVSQSSAVNISMQAWSLWPKRKQRAFFLFGLELIVOLDIEATRTFFTRFLPT 474
Db 416 KOINS-----NISQAWDTLWPPERKQRAFFLFGLELIVQFDEGIRSFTRFLPK 469
QY 475 WNWGFLGSSLSFDLVLFSTMTFVLAPNSMRSLVRHLLSDPSGAVMVRAYLE 528

Db 416 KOINS-----NISQAWDTLWPPERKQRAFFLFGLELIVQFDEGIRSFTRFLPK 469
QY 475 WNWGFLGSSLSFDLVLFSTMTFVLAPNSMRSLVRHLLSDPSGAVMVRAYLE 528
Db 470 WNWGFLGSLTSGDLVLFALYMFVISPNNLRKGLNHLISDPTGATMIKTYLK 523

RESULT 7
US-09-323-998D-49
; Sequence 49, Application US/09323998D
; Patent No. US20020102631A1
; GENERAL INFORMATION:
; APPLICANT: CUNNINGHAM JR., FRANCIS X.
; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND
; FILE REFERENCE: 108172-09019
; CURRENT APPLICATION NUMBER: US/09/323,998D
; PRIOR FILING DATE: 1999-06-02
; PRIOR FILING DATE: 1998-06-02
; PRIOR FILING DATE: 1998-06-02
; PRIOR FILING DATE: 1997-09-25
; PRIOR FILING DATE: 1996-03-29
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 49
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-323-998D-49

Query Match 71.4%; Score 1959.5; DB 10; Length 524;
Best Local Similarity 70.4%; Pred. No. 1.9e-174;
Matches 376; Conservative 60; Mismatches 81; Indels 17; Gaps 5;

QY 1 MELLGVRNL-----ISSCPVMTFTGRNLSKSLAYNIHRYGSSCRVDFQVRADGG--SGSR 54
Db 1 MECVGARNFAMAVSTPFSKCRKFPVKRYISYRNIRFG--LC-----SVRASGGSSGE 55
QY 55 SSVAYKEGFVDEEDFIKAGSELFFVQMOKTSMKQAKLADKLPIPPFGESVMDLVVIG 114
Db 56 SCVAVREDFADEEDFVKAGSEILFFVQMOKMDQSKLVKLPPIISIGDGLDHWVIG 115
QY 115 CGPAGLSLAAEAAKGLKVLGDLPTFTNNYGVWDEDFDKGLERCIEHAWKDTIVYLD 174
Db 116 CGPAGLALAAESAKGLKVLGDLPTFTNNYGVWDEDFDKGLERCIEHAWKDTIVYLD 175
QY 175 NDAPVLITGRAYGRVSRHLLHEELLKRCVSGSVYLSKVERITEAGDGHSLVVCENEIFI 234
Db 176 DDKPITIGRAYGRVSRHLLHEELLKRCVSGSVYLSKVDISITEASDGLRLVACDNNVI 235
QY 235 PCRLATVASGAAGSKLLEYEVGGPRVCVQYATGYVEVEVNNPYDPNLMVPMYRDYMOOK 294
Db 236 PCRLATVASGAAGSKLLEYEVGGPRVCVQYATGYVEVEVNNPYDPNLMVPMYRDYMOOK 295
QY 295 LOCSEEEYPTFLVMPMSPTFLFEETCTCLASKDAMPFDLLKRLKMSRLKTLGTVKTYE 354
Db 296 VRSLEAEYPTFLVMPMSPTFLFEETCTCLASKDAMPFDLLKRLKMSRLKTLGTVKTYE 355
QY 355 EWSYIPVGGSLPNTQKNLAFGAAASWHPATGYSVRSLSSEAPKYASVIAEILREET 414
Db 356 EWSYIPVGGSLPNTQKNLAFGAAASWHPATGYSVRSLSSEAPKYASVIAEILREET 415
QY 415 AYVVSQSSAVNISMQAWSLWPKRKQRAFFLFGLELIVOLDIEATRTFFTRFLPT 474
Db 416 KOINS-----NISQAWDTLWPPERKQRAFFLFGLELIVQFDEGIRSFTRFLPK 469
QY 475 WNWGFLGSSLSFDLVLFSTMTFVLAPNSMRSLVRHLLSDPSGAVMVRAYLE 528

Db 470 NMWQFLGSLTSLTSGDLVLFALYMFVISPNNLRKGLNHLISDPTGATMIKTYLK 523

RESULT 8

US-09-323-998D-54
; Sequence 54, Application US/09323998D
; Patent No. US20020102631A1
; GENERAL INFORMATION:
; APPLICANT: CUNNINGHAM JR., FRANCIS X.
; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND
; FILE REFERENCE: 108172-09019
; CURRENT FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: US/09/323,998D
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 09/088,724
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 09/088,725
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 08/937,155
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 08/624,125
; PRIOR FILING DATE: 1996-03-29
; SOFTWARE: Patent In Ver. 2.1
; NUMBER OF SEQ ID NOS: 61
; SEQ ID NO 54
; LENGTH: 516
; TYPE: PRM
; ORGANISM: Tagetes erecta
US-09-323-998D-54

Query Match 70.5%; Score 1934; DB 10; Length 516;
Best Local Similarity 71.4%; Pred. No. 4,4e-172;
Matches 365; Conservative 64; Mismatches 74; Indels 8; Gaps 2;
QY 22 NLSSSKLAYNTHRYGSSCRVDFQVRADGGSSRSRVAYKEGVDFEEDFKAGGSELLFVQ 81
Db 7 HMTATMAAFTCPREMTSIRYTKQIKC---NAKSQLVVKQETEEEDYVKAGGSELLFVQ 63
QY 82 MOOTKSMKQAKLADKLPPIPG-----ESVMDLVVIGCGPAGLSIAEAAKGLKVGLI 136
Db 64 MOONKSMDAQSLSQKLPVPVIGGGSDNCILDLVVIGCGPAGLALAGESAKGLNVALI 123
QY 137 GPDLPFTNNGYWEDEFKDLGLERIEHAWKDTIYLDNDAPVLIGRAYGRVSRHLLHEE 196
Db 124 GPDLPFTNNGYWEDEFKDLGLERIEHAWKDTIYLDNDAPVLIGRAYGRVSRHLLHEE 183
QY 137 LLKRCVESGVYLDKSKVERITEAGDCHSLVVCENEIFIPCRLATVAGSAAGKLLLEYEYV 256
Db 184 LLTRCMESGVYLDKSKVERITEAGDCHSLVVCENEIFIPCRLATVAGSAAGKLLLEYEYV 243
QY 257 GPRVCVQYAYGEVEVENNPDPNLMVMDYRDYMQKLCQSEEEYPTFLYVMPSPTRL 316
Db 244 GPRVCVQYAYGEVEVENNPDPNLMVMDYRDYMQKLCQSEEEYPTFLYVMPSPTKV 303
QY 317 FFEETCLASKAMPEDLLKRLMSRLKTLGIQVTKVYEEENSYIPVGGSLPNTQKNLAF 376
Db 304 FFEETCLASKAMPEDLLKRLMSRLKTLGIQVTKVYEEENSYIPVGGSLPNTQKNLAF 363
QY 377 GAAASWHPATGYSVVRSLSEAPKVASYIAKTLKODNSAYVVGSSAVNISQAWSSLSW 436
Db 364 GAAASWHPATGYSVVRSLSEAPKVASYIAKTLKODNSAYVVGSSAVNISQAWSSLSW 423
QY 437 PKERKQRAFFLFGLELIVQMDIEATRTFFRFLPTMWMWGLGSSLSFDDLVLFMSY 496
Db 424 PKERKQRAFFLFGLELIVQMDIEATRTFFRFLPTMWMWGLGSSLSFDDLVLFMSY 483
QY 497 MFVLAPNSMRSLVRHLLSDPSGAVMVRAYL 527
Db 484 MFVLAPNSMRSLVRHLLSDPSGAVMVRAYL 514

RESULT 9

US-09-323-998D-52
; Sequence 52, Application US/09323998D
; Patent No. US20020102631A1
; GENERAL INFORMATION:
; APPLICANT: CUNNINGHAM JR., FRANCIS X.
; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND
; FILE REFERENCE: 108172-09019
; CURRENT FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: US/09/323,998D
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 09/088,724
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 09/088,725
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 08/937,155
; PRIOR FILING DATE: 1997-09-25
; PRIOR APPLICATION NUMBER: 08/624,125
; PRIOR FILING DATE: 1996-03-29
; SOFTWARE: Patent In Ver. 2.1
; NUMBER OF SEQ ID NOS: 61
; SEQ ID NO 52
; LENGTH: 533
; TYPE: PRM
; ORGANISM: Lactuca sp.
US-09-323-998D-52

Query Match 70.5%; Score 1934; DB 10; Length 533;
Best Local Similarity 68.3%; Pred. No. 4,7e-172;
Matches 375; Conservative 57; Mismatches 77; Indels 40; Gaps 5;
QY 1 MELLIGVRNI-----TSSCPVWT-----FGRNLSSSKLAYNTHRYGSSCR 40
Db 1 MCFGARNMTATMAVETCPRTDTCNTRHKFSLKQRRFTNLSASSLSURQIKCSAKSDRCV 60
QY 41 VDFQVRADGGSSRSRVAYKEGF--VDEEDFKAGGSELLFVQMOOTKSMKQAKLADK 98
Db 61 VD-----KQGISVADEEDYVKAGGSELLFVQMOOTKSMKQAKLADK 103
QY 99 PPIPGESVMDLVVIGCGPAGLSIAEAAKGLKVGLIGPDLPFTNNGYWEDEFKDLGL 158
Db 104 AQIPGNCILDLVVIGCGPAGLALAGESAKGLNVALI 163
QY 159 ERCIEHAWKDTIYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVESGVYLDKSKVERITE 218
Db 164 ERCIEHAWKDTIYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVESGVYLDKSKVERITE 223
QY 219 AGDGHSLVVCENEIFIPCRLATVAGSAAGKLLLEYEYVGGPRVCVQYAYGEVEVENNPYD 278
Db 224 APNGYSLIECEGNITIPCRLATVAGSAAGKLLLEYEYVGGPRVCVQYAYGEVEVENNPYD 283
QY 279 PNLVFMVMDYRDYMQKLCQSEEEYPTFLYVMPSPTRLFFETCLASKAMPEDLLKRL 338
Db 284 PNLVFMVMDYRDYMQKLCQSEEEYPTFLYVMPSPTRLFFETCLASKAMPEDLLKRL 343
QY 339 MSRLTGLIQVTKVYEEENSYIPVGGSLPNTQKNLAFGAAASWHPATGYSVVRSLSEA 398
Db 344 MSRLKAMGIRIYRTYEEENSYIPVGGSLPNTQKNLAFGAAASWHPATGYSVVRSLSEA 403
QY 399 PKYASVIKILKODNSAYVVGSSAVNISQAWSSLSWPKERKQRAFFLFGLELIVQMD 458
Db 404 PNYAAVIAKILKODNSAYVVGSSAVNISQAWSSLSWPKERKQRAFFLFGLELIVQMD 462
QY 459 ICATRTFFRFLPTMWMWGLGSSLSFDDLVLFMSYMFVLAPNSMRSLVRHLLSDPS 518
Db 463 LEGTTFTRFFRFLPKMWMWGLGSSLSFDDLVLFMSYMFVLAPNSMRSLVRHLLSDPS 522
QY 519 GAVMVRAYL 527
Db 523 GATMVRAYL 531

RESULT 10

US-09-323-998D-53
; Sequence 53, Application US/09323998D
; Patent No. US20020102631A1
; GENERAL INFORMATION:
; APPLICANT: CUNNINGHAM JR., FRANCIS X.
; APPLICANT: SUN, ZAIREN
; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 108172-09019
; CURRENT APPLICATION NUMBER: US/09/323,998D
; CURRENT FILING DATE: 1999-06-02
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: 09/088,724
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 09/088,725
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 08/937,155
; PRIOR FILING DATE: 1997-09-25
; PRIOR APPLICATION NUMBER: 08/624,125
; PRIOR FILING DATE: 1996-03-29
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 53
; LENGTH: 526
; TYPE: PRT
; ORGANISM: Lycopodium esculentum
US-09-323-998D-53

Query Match 69.6%; Score 1909.5; DB 10; Length 526;
Best Local Similarity 68.5%; Pred. No. 8.9e-170;
Matches 368; Conservative 68; Mismatches 78; Indels 23; Gaps 6;
QY 1 MELLGVNRLISSCPV-----WTFGRNLSKK---LAYNIHRYGSSCRVDFOVRADGG 50
DB 1 MECVGQVNGVAMAYLTPRLNRWGG--ELCQKSIIFL--EQYESKC-----NSS 48
QY 51 SGRSSVAYKEGFDEDFIKAGSELFFVQMOTKMEKQAKLADKLPIPPGSESVDL 110
DB 49 SGRSSCVVDKEDFDEDFIKAGSELFFVQMOTKMEKQAKLADKLPIPPGSESVDL 108
QY 111 VVIGCGPAGLSLAAEAAKGLKVLGDPDLPFTNNYGVWEDEFKDLGLERIEHAWKDTI 170
DB 109 VVIGCGPAGLSLAAEAAKGLKVLGDPDLPFTNNYGVWEDEFKDLGLERIEHAWKDTI 168
QY 171 VYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVSGVSYLDSKVERITEAGDGLSVVCEN 230
DB 169 VYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVSGVSYLDSKVERITEAGDGLSVVCEN 228
QY 231 EIFTPCRLATVAGSAAGSKLLEYEVGPRVCQTAYGVEVEVNNPYDNLVFMFMDYRDY 290
DB 229 DVVTPCRFVTVAGSAAGSKLLEYEVGPRVCQTAYGVEVEVNNPYDNLVFMFMDYRDY 288
QY 291 MQOKLQCSSEYPTFLVMPMSPTLFFETCLASKDAMPFDLLKRLKSLKTLGIQVT 350
DB 289 LRHDAQSLKATPTFLVMPMSPTLFFETCLASKDAMPFDLLKRLKSLKTLGIQVT 348
QY 351 KYVEEWSYIPVGSGLPNTQKNLAFGAAASMVHPATGYSVRSLSSEAPYASVIKILK 410
DB 349 EYVEEWSYIPVGSGLPNTQKNLAFGAAASMVHPATGYSVRSLSSEAPYASVIKILK 408
QY 411 QDNVYVSGSSAVNISMQAWSLWPKERKQRAFFLGLGLTIVQDIEATFFFTFF 470
DB 409 QHYSKNMLT--SSSTPSTQAWNTLWPKERKQRAFFLGLGLTIVQDIEATFFFTFF 467
QY 471 RLPTWMMWGLGSSLSFDLVLFVSMYFVLAPNSMRSLVRLHLLSDPSPGAVMVRAYL 527
DB 468 RVPKMMWQGLGSSLSFDLVLFVSMYFVLAPNSMRSLVRLHLLSDPSPGAVMVRAYL 524

RESULT 11
US-09-323-998D-26
; Sequence 26, Application US/09323998D
; Patent No. US20020102631A1
; GENERAL INFORMATION:

; APPLICANT: CUNNINGHAM JR., FRANCIS X.
; APPLICANT: SUN, ZAIREN
; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 108172-09019
; CURRENT APPLICATION NUMBER: US/09/323,998D
; CURRENT FILING DATE: 1999-06-02
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: 09/088,724
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 09/088,725
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 08/937,155
; PRIOR FILING DATE: 1997-09-25
; PRIOR APPLICATION NUMBER: 08/624,125
; PRIOR FILING DATE: 1996-03-29
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 26
; LENGTH: 533
; TYPE: PRT
; ORGANISM: Artificial Sequence
; OTHER INFORMATION: Description of Artificial Sequence: Lactuca sp./Solanum sp.
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (491)
; OTHER INFORMATION: Any amino acid
US-09-323-998D-26

Query Match 68.6%; Score 1882; DB 10; Length 533;
Best Local Similarity 67.6%; Pred. No. 3.4e-167;
Matches 370; Conservative 65; Mismatches 76; Indels 36; Gaps 8;
QY 1 MELLGVNRL-----ISSCPVWT-----FGTR--NLSSSKLAYNIHRYGSSCRVD 42
DB 1 MECVGARNMATMAVFTCPRTDCNIRHFKSLKGRRTNLSAS-----SSLR-- 48
QY 43 FQVRADGSGSSRSVAYKEGF--VDEEDFIKAGSELFFVQMOTKMEKQAKLADKLPP 100
DB 49 -QIKC--SAKSDRCVVDKQIGSVDAEDYKAGSELFFVQMOTKMEKQAKLADKLPP 105
QY 101 IPGSEVMDLVVIGCGPAGLSLAAEAAKGLKVLGDPDLPFTNNYGVWEDEFKDLGLER 160
DB 106 IPGSEVMDLVVIGCGPAGLSLAAEAAKGLKVLGDPDLPFTNNYGVWEDEFKDLGLER 165
QY 161 CIEHAWKDTIVYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVSGVSYLDSKVERITEAG 220
DB 166 CIEHAWKDTIVYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVSGVSYLDSKVERITEAG 225
QY 221 DGHSLVVCENEIETPCRLATVAGSAAGSKLLEYEVGPRVCQTAYGVEVEVNNPYDNP 280
DB 226 NGYSLICEGNIETPCRLATVAGSAAGSKLLEYEVGPRVCQTAYGVEVEVNNPYDNP 285
QY 281 LAMFMDYRDYMQOKLQCSSEYPTFLVMPMSPTLFFETCLASKDAMPFDLLKRLKLS 340
DB 286 LAMFMDYRDYMQOKLQCSSEYPTFLVMPMSPTLFFETCLASKDAMPFDLLKRLKLS 345
QY 341 RLKTLGIQVTKVYEEWSYIPVGSGLPNTQKNLAFGAAASMVHPATGYSVRSLSSEAPK 400
DB 346 RLKTLGIQVTKVYEEWSYIPVGSGLPNTQKNLAFGAAASMVHPATGYSVRSLSSEAPK 405
QY 401 YASVIKILKQDNVYVSGSSAVNISMQAWSLWPKERKQRAFFLGLGLTIVQDIE 460
DB 406 CAFVLANILQDNVYVSGSSAVNISMQAWSLWPKERKQRAFFLGLGLTIVQDIE 464
QY 461 ATFTFTFTFRLPTWMMWGLGSSLSFDLVLFVSMYFVLAPNSMRSLVRLHLLSDPSPA 520
DB 465 GRSFTRAFTRVVPKMMWQGLGSSLSFDLVLFVSMYFVLAPNSMRSLVRLHLLSDPSPA 524
QY 521 VMVRAYL 527
DB 525 TLIRTYL 531

RESULT 12
US-09-323-998D-25
; Sequence 25, Application US/09323998D
; Patent No. US20020102631A1
; GENERAL INFORMATION:
; APPLICANT: CUNNINGHAM JR., FRANCIS X.
; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND
; FILE REFERENCE: 108172-09019
; CURRENT APPLICATION NUMBER: US/09/323,998D
; CURRENT FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: 09/088,724
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 09/088,725
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 08/937,155
; PRIOR FILING DATE: 1997-09-25
; PRIOR APPLICATION NUMBER: 08/624,125
; PRIOR FILING DATE: 1996-03-29
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 25
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Solanum tuberosum
; FEATURE:
; NAME/KEY: MOD.RES
; LOCATION: (336)
; OTHER INFORMATION: Any amino acid
US-09-323-998D-25

Query Match 55.4%; Score 1518.5; DB 10; Length 378;
Best Local Similarity 74.5%; Pred. No. 1.9e-133;
Matches 281; Conservative 48; Mismatches 47; Indels 1; Gaps 1;
QY 151 DEFKDLGERCTEHAWKDTIVYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVSGVSYLD 210
Db 1 DEFKDLGLQACIEHVWRDTIVYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVSGVSYLD 60
QY 211 SKVERITEAGDCHSLVVCENEIFPCRLATVASGAASKLLEYEYVGGPRVCVQTAYGVEV 270
Db 61 SKVDRIEATNGHSLVCEGDDVPCRFVTVASGAASKGLQYELGPRVSVQTAYGVEV 120
QY 271 EVNNPYDNLVMDYRDYMOQKQCSEEEYPTFLYVMPSPTRLFFETCLASKDAMP 330
Db 121 EVDNPFDSLVMFYMDYRDYVHDAQSLQAKYPTFLYAMPSPTRVFFETCLASKDAMP 180
QY 331 FDLKRLKMSRLKLTIGIQVTKYEEWSYIPVGGSLPNTQKNLAFGAASVHPATGYS 390
Db 181 FDLKRLKMLRLNTLGVRIKEIYEWSYIPVGGSLPNTQKTLAFGAASVHPATGYS 240
QY 391 VVRSLEAPKVASVTAIILKQNSAVVYVSGSSAVNISQAWSSLPKPKRQRAFFLFG 450
Db 241 VVRSLEAPKCAFVLANTLRQNHKNMLT-SSSTPSTQAWNTLWPKRQRAFFLFG 299
QY 451 LELIVQLDIEATRTFFRFTFRFTFRFTFRFTFRFTFRFTFRFTFRFTFRFTFRFTFR 510
Db 300 LALIQLDIEGIRSFRAFFRVKPMWQGLGSSLSXADLMFLAFYMFIIAPNDMRRGLI 359
QY 511 RHLLSDPSGAVMVRAYL 527
Db 360 RHLLSDPTGATLIRTYL 376

RESULT 13
US-09-323-998D-48
; Sequence 48, Application US/09323998D
; Patent No. US20020102631A1
; GENERAL INFORMATION:
; APPLICANT: CUNNINGHAM JR., FRANCIS X.
; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND

; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND
; FILE REFERENCE: 108172-09019
; CURRENT APPLICATION NUMBER: US/09/323,998D
; CURRENT FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: 09/088,724
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 09/088,725
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 08/937,155
; PRIOR FILING DATE: 1997-09-25
; PRIOR APPLICATION NUMBER: 08/624,125
; PRIOR FILING DATE: 1996-03-29
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 48
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Solanum tuberosum
; FEATURE:
; NAME/KEY: MOD.RES
; LOCATION: (336)
; OTHER INFORMATION: Any amino acid
US-09-323-998D-48

Query Match 55.4%; Score 1518.5; DB 10; Length 378;
Best Local Similarity 74.5%; Pred. No. 1.9e-133;
Matches 281; Conservative 48; Mismatches 47; Indels 1; Gaps 1;
QY 151 DEFKDLGERCTEHAWKDTIVYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVSGVSYLD 210
Db 1 DEFKDLGLQACIEHVWRDTIVYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVSGVSYLD 60
QY 211 SKVERITEAGDCHSLVVCENEIFPCRLATVASGAASKLLEYEYVGGPRVCVQTAYGVEV 270
Db 61 SKVDRIEATNGHSLVCEGDDVPCRFVTVASGAASKGLQYELGPRVSVQTAYGVEV 120
QY 271 EVNNPYDNLVMDYRDYMOQKQCSEEEYPTFLYVMPSPTRLFFETCLASKDAMP 330
Db 121 EVDNPFDSLVMFYMDYRDYVHDAQSLQAKYPTFLYAMPSPTRVFFETCLASKDAMP 180
QY 331 FDLKRLKMSRLKLTIGIQVTKYEEWSYIPVGGSLPNTQKNLAFGAASVHPATGYS 390
Db 181 FDLKRLKMLRLNTLGVRIKEIYEWSYIPVGGSLPNTQKTLAFGAASVHPATGYS 240
QY 391 VVRSLEAPKVASVTAIILKQNSAVVYVSGSSAVNISQAWSSLPKPKRQRAFFLFG 450
Db 241 VVRSLEAPKCAFVLANTLRQNHKNMLT-SSSTPSTQAWNTLWPKRQRAFFLFG 299
QY 451 LELIVQLDIEATRTFFRFTFRFTFRFTFRFTFRFTFRFTFRFTFRFTFRFTFRFTFR 510
Db 300 LALIQLDIEGIRSFRAFFRVKPMWQGLGSSLSXADLMFLAFYMFIIAPNDMRRGLI 359
QY 511 RHLLSDPSGAVMVRAYL 527
Db 360 RHLLSDPTGATLIRTYL 376

RESULT 14
US-09-323-998D-27
; Sequence 27, Application US/09323998D
; Patent No. US20020102631A1
; GENERAL INFORMATION:
; APPLICANT: CUNNINGHAM JR., FRANCIS X.
; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND
; FILE REFERENCE: 108172-09019
; CURRENT APPLICATION NUMBER: US/09/323,998D
; CURRENT FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: 09/088,724
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 09/088,725

; PRIOR FILING DATE: 1998-06-02
 ; PRIOR APPLICATION NUMBER: 08/937,155
 ; PRIOR FILING DATE: 1997-09-25
 ; PRIOR APPLICATION NUMBER: 08/624,125
 ; PRIOR FILING DATE: 1996-03-29
 ; NUMBER OF SEQ ID NOS: 61
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO: 27
 ; LENGTH: 374
 ; TYPE: PRT
 ; ORGANISM: Arabidopsis thaliana
 ; US-09-323-998D-27

Query Match 55.2%; Score 1513; DB 10; Length 374;
 Best Local Similarity 74.7%; Pred. No. 6e-133;
 Matches 283; Conservative 42; Mismatches 48; Indels 6; Gaps 1;
 QY 150 EDEFKDLGERCIHAWKDTIIVLDNDAPVLIGRAYGRVSRHLLHELLKRCVSGSYL 209
 Db 1 EDEFNDLGLQKCIHVMRETIVYLDKPTIGRAYGRVSRHLLHELLKRCVSGSYL 60
 QY 210 DSKVERITEAGDGHSLVVCENEIFIPCRLATVASGAAGKLLLEVGGPRVCQYAGVE 269
 Db 61 SSRVDSITEASDGLRLVACDNNVPCRLATVASGAAGKLLQYEVGGPRVCQYAGVE 120
 QY 270 VEENNPYDNLVMDYRDYMOQKLCQSEEEYPTLYVMPSPTRLFFETCTCLASKDAM 329
 Db 121 VEENSPYDQVMYRDYTNKVRSLAEAYPTLYAMPMTKSLRFFETCTCLASKDYM 180
 QY 330 PFLLKRLKRLKTLGIGQVTKVEEWSYIPVGGSLPNTPEQNLAFGAASVMVHPATGY 389
 Db 181 PFLLKRLKRLKTLGIRLTKYTEEWSYIPVGGSLPNTPEQNLAFGAASVMVHPATGY 240
 QY 390 SVYRSLSEAPKYASVIAKILKQDNASVYVSGSSAVNISQMSLWPKRKQRAFFELF 449
 Db 241 SVYRSLSEAPKYASVIAELRETTQKINS-----NISQAWDTLWPKRKQRAFFELF 294
 QY 450 GLELIVOLDIEATRTFFRFLPTMWMGFLGSSLSFSLVLFMSYMFVLAAPNSMRMSL 509
 Db 295 GLALIVQDFEGIRSFERTFRLPKMWMQGLGSLTSLTSGDLVLFALYMFVISPNNLRKGL 354
 QY 510 VRHLSDPSCAVMVRAYLE 528
 Db 355 INHLSDPTGATMIKTYLK 373

RESULT 15
 US-09-323-998D-56
 ; Sequence 56, Application US/09323998D
 ; Patent No. US20020102631A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CUNNINGHAM JR., FRANCIS X.
 ; APPLICANT: SUN, ZAIREN
 ; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND METABOLISM AND
 ; FILE OF INVENTION: METHODS OF USE THEREOF
 ; FILE REFERENCE: 108172-09019
 ; CURRENT APPLICATION NUMBER: US/09/323,998D
 ; CURRENT FILING DATE: 1999-06-02
 ; PRIOR APPLICATION NUMBER: 09/088,724
 ; PRIOR FILING DATE: 1998-06-02
 ; PRIOR APPLICATION NUMBER: 09/088,725
 ; PRIOR FILING DATE: 1998-06-02
 ; PRIOR APPLICATION NUMBER: 08/937,155
 ; PRIOR FILING DATE: 1997-09-25
 ; PRIOR APPLICATION NUMBER: 08/624,125
 ; PRIOR FILING DATE: 1996-03-29
 ; NUMBER OF SEQ ID NOS: 61
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 56
 ; LENGTH: 502
 ; TYPE: PRT
 ; ORGANISM: Adonis palaestina
 ; US-09-323-998D-56

Query Match 28.3%; Score 775.5; DB 10; Length 502;
 Best Local Similarity 36.3%; Pred. No. 8e-64;
 Matches 181; Conservative 99; Mismatches 166; Indels 53; Gaps 14;
 QY 23 LSSSKLAYNIHRYGS-----SCRVDFOVRADGGSGSRSSVAYKEGFVDEEDFIKAGGSELL 78
 Db 27 VSTSKLQNOVFRIASRNHPCR-----NGT-----VKARGSALL 60
 QY 79 FVQMOQTFSMEKQAKLADKLPPPI-PGESVMDLVVTCGCGPAGLSLAAEAANKLGLVGLI- 136
 Db 61 -----ELVPETKKNLEFDLPAYDPSRGIVVDLAVVGGGPGAGLAIQAQVSEAGLLVCSID 115
 QY 137 -GPDLPFTNNYGVWEDEFKDLGERCIHAWKDTIIVLDNDAPVLIGRAYGRVSRHLLHE 195
 Db 116 PPSKLLWPNNYGVWDEFEAMDLLDCLDTTWSGAVVYTDNKKYLDLDRPYGRVNRKQLKS 175
 QY 196 ELKRCVSGSVYLDKSKVERITEAGDGHSLVVCENEIFIPCRLATVASGAAGKLLLEYEV 255
 Db 176 KMLQKCVTNGVKPHQAKVIKVIHE-ESKSLICNDGTTINATVVLDTATG-FSRCLVQYDK 233
 QY 256 G-GPRVCVQYAGVEVEENNPYDNLVMDYRD-YMOOKLQCSSE--EYPTFLYVMPM 311
 Db 234 PYNPG--YQVAYGIMAEVEEHPFDLKMFLMDWRDLSHLNEKLEKDKNRKIPITFLYAMPF 291
 QY 312 SPTRLEFEETCLASKDAMPFDLKRKLMSRLKTLGIGQVTKVEEWSYIPVGGSLPNTPEQ 371
 Db 292 SSKPILEETSLVARPGLRFEDIOERVARLKHGKVKSIEDERCVPIMGGPLPVLPO 351
 QY 372 KNLAFGAASVMVHPATGYSVYRSLSEAPKYASVIAKILKQDNASVYVSGSSAVNISQMA 431
 Db 352 RVVGIGGTAGMVHPSTGYMVARTLAAAPVVAKSIVQYLGSDRS---LSGN---ELSAEV 404
 QY 432 WSLWPKRKQRAFFELGLELIVOLDIEATRTFFRFLPTMWMGFLGSSLSFSLV 491
 Db 405 WDLWPIERRRQREFFCFGMDILLKLDLQGTTRFFDAFFDLEPHYWHGFLSLRFLPELL 464
 QY 492 LFSMYMFLVAPNSMRMSLV 510
 Db 465 FFLGLSLFASHASNASRIEM 483

Search completed: May 21, 2003, 22:22:05
 Job time : 192 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 21, 2003, 21:30:47 ; Search time 22 Seconds
(without alignments)
997.317 Million cell updates/sec

Title: US-09-701-395A-23

Perfect score: 2743

Sequence: 1 MELLGVRNLISCPVWTFGT.....VRHLLDPGAVMVRAYLER 529

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1959.5	71.4	524	1 LCYE_ARATH	Q38932 arabidopsis
2	1909.5	69.6	526	1 LCYE_LYCES	O65837 lycopersico
3	761	27.7	498	1 LCYE_CAPAN	Q43415 capsicum an
4	749.5	27.3	500	1 LCYE_LYCES	Q43503 lycopersico
5	748	27.3	501	1 LCYE_ARATH	Q38933 arabidopsis
6	737	26.9	500	1 LCYE_TOBAC	Q43578 nicotiana t
7	708	25.8	503	1 CCS_CITSI	Q9sea0 citrus sine
8	705.5	25.7	498	1 CCS_CAPAN	Q42435 capsicum an
9	701.5	25.6	503	1 LCYE_NARES	Q49424 narcissus p
10	602	21.9	411	1 LCYE_SYNP7	Q52726 synechococc
11	429.5	15.7	410	1 Y801_DEIRA	Q9rw68 deinococcus
12	137.5	5.0	506	1 ER11_BRANA	O65727 brassica na
13	129	4.7	386	1 CRTY_AGRAU	P54974 agrobacteri
14	125.5	4.6	382	1 CRTY_PANAN	P21687 pantoea ana
15	125	4.6	386	1 CRTY_ERWHE	Q01331 erwinia her
16	112	4.1	518	1 ER12_BRANA	O65726 brassica na
17	111	4.0	517	1 ER12_ARATH	O65402 arabidopsis
18	109.5	4.0	391	1 Y532_METUA	O57952 methanococc
19	106.5	3.9	372	1 FTZ1_PYRAB	O9v280 pyrococcus
20	105.5	3.8	414	1 YD00_SYNY3	P72835 synechocyst
21	105	3.8	539	1 ERG1_PANGI	O48651 panax ginsu
22	105	3.8	593	1 Y40A_RHISN	P55586 rhizobium s
23	104.5	3.8	516	1 ER13_ARATH	O65403 arabidopsis
24	103	3.8	472	1 PCNB_ECOLI	P13685 escherichia
25	103	3.8	510	1 C4DK_DROME	O9w011 drosophila
26	101.5	3.7	372	1 FTZ1_PYRHO	O57776 pyrococcus
27	101.5	3.7	556	1 FTHS_STRMO	Q59925 streptococc
28	100.5	3.7	366	1 FTZ1_PYRWO	Q52630 pyrococcus
29	100.5	3.7	497	1 TRXB_HUMAN	Q16881 homo sapien
30	100	3.6	1071	1 CARB_BACSU	P25994 bacillus su
31	100	3.6	1202	1 ALAA_ARATH	Q91183 arabidopsis
32	99.5	3.6	308	1 R34K_CLOPA	P23160 clostridium
33	99.5	3.6	1064	1 CARB_LACLC	O32771 lactococcus

ALIGNMENTS

RESULT 1

ID	LCYE_ARATH	STANDARD;	PRT;	524 AA.
AC	Q38932;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Lycopene epsilon cyclase, chloroplast precursor (EC 1.14.-.-).			
OS	Arabidopsis thaliana (Mouse-ear cress).			
OC	Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;			
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsis.			
OX	NCBI_Taxid=3702;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=cv. Columbia;			
RX	MEDLINE=96434545; PubMed=8837512;			
RA	Cunningham F.X. Jr., Pogson B., Sun Z., McDonald K.A., Dellapenna D.,			
RA	Gantt E.;			
RT	*Functional analysis of the beta and epsilon lycopene cyclase enzymes			
RT	of Arabidopsis reveals a mechanism for control of cyclic carotenoid			
RT	formation.*;			
RL	Plant Cell 8:1613-1626(1996).			
CC	- FUNCTION: CATALYZES THE SINGLE CYCLIZATION REACTION WHICH CONVERTS			
CC	LYCOPENE TO DELTA-CAROTENE AND NEUROSPORENE TO ALPHA-ZEACAROTENE.			
CC	- PATHWAY: Carotenoid biosynthesis.			
CC	- SUBCELLULAR LOCATION: Chloroplast.			
CC	- SIMILARITY: BELONGS TO THE LYCOPENE CYCLASE FAMILY.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
DR	EMBL; U50738; AA853336.1; -			
KW	Oxidoreductase; NAD; Carotenoid biosynthesis; Chloroplast;			
FT	Transit peptide.			
FT	TRANSIT 1 ? CHLOROPLAST (POTENTIAL).			
FT	CHAIN ? 524 LYCOPENE EPSILON CYCLASE.			
FT	NP_BIND 111 139 NAD (POTENTIAL).			
SQ	SEQUENCE 524 AA; 58515 MW; F4A4BBC9D5487E1 CRC64;			
Query Match 71.4%; Score 1959.5; DB 1; Length 524;				
Best Local Similarity 70.4%; Pred No. 7.9e-143;				
Matches 376; Conservative 60; Mismatches 81; Indels 17; Gaps 5;				
QY	1 MELLGVRNL-----ISSCPVWTFGTNRNLSSKSLAYNIHRYGSSCRVDQVRADGG--SGSR 54			
	: : : : : : : : : : : :			
Db	1 MECVGARNEFAAMAVSTFPSCRRKFPVKYRYRNIREG-LC-----SVRASGGSGSGS 55			
QY	55 SSVAKEGFVDEEDFTKAGSGELLFVQMOTKSMKQAKLADKLPIPIPGESVMDLWVG 114			
	: : : : : : : : : : :			
Db	56 SCVAVREDFADEEDFYKAGSGELLFVQMOTKSMKQAKLADKLPIPIPGESVMDLWVG 115			

Query Match 69.6%; Score 1909.5; DB 1; Length 526;
 Best Local Similarity 68.5%; Pred. No. 5.5e-139;
 Matches 369; Conservative 68; Mismatches 78; Indels 23; Gaps 6;

QY 115 CGPAGLSLAALAAKGLKVLGPDLPFTNNYGVWDEDFKDLGLERCIEHAWKDTIIVYLD 174
 D 116 CGPAGLALAAESAKGLKVLGPDLPFTNNYGVWDEDFKDLGLERCIEHAWKDTIIVYLD 175
 QY 175 NDAPVLIGRAYGRVSRHLLHEELKRCVSGSVLDSKVERITEAGDGHSLVVCENEIFI 234
 D 176 DDKPITIGRAYGRVSRHLLHEELKRCVSGSVLDSKVERITEAGDGHSLVVCENEIFI 235
 QY 235 PCRLATVASGAASGLLEYEVGGPRVCQYAYGVEVEVNNPYDPNLMVMDYRDYMOOK 294
 D 236 PCRLATVASGAASGLLQYEVGGPRVCQYAYGVEVEVNSPYDPQVMFYMDYRDYNEK 295
 QY 295 LQSESEYPTFLYVMPSPRLFFETCLASKDAMPDLKRLKMSRLKTLGIQVTKYVE 354
 D 296 VRSLEAEYPTFLYVMPSPRLFFETCLASKDAMPDLKRLKMSRLKTLGIQVTKYVE 355
 QY 355 EEWYIPVGGSLPNTQKNLAFGAASWVHPATGYSVVRSLSEAPKASYIAKILKQDNS 414
 D 356 EEWYIPVGGSLPNTQKNLAFGAASWVHPATGYSVVRSLSEAPKASYIAKILRETT 415
 QY 415 AYVVGSSAVNISMQAWSSLPKRRQRRAFFLGLGLIVQLDIEATRTFFFRFLPT 474
 D 416 KQINS-----NISQAWDTLWPPERKRRQRAFFLGLGLIVQLDIEATRTFFFRFLPK 469
 QY 475 WMMWGLGSLSSDPLVFSYMFVLPAPNSMRSLVRHLLSDPGAVVWVAYLE 528
 D 470 WMMWGLGSLTSGDLVLFALYMEVISPNNLRKGLNHLISDPTGATWIKTYLK 523

RESULT 2

LCYE_LICES STANDARD; PRT; 526 AA;
 AC 065837;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Lycopen epsilon cyclase, chloroplast precursor (EC 1.14.-.-).
 GN CRTL-E-1.
 OS Lycopersicon esculentum (Tomato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=4081;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. VF36; TISSUE=Leaf;
 RA Ronen G., Cohen M., Zamir D., Hirschberg J.;
 RT "Regulation of expression of the gene for lycopen epsilon cyclase
 during fruit ripening of tomato.";
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 CC -|- FUNCTION: CATALYZES THE SINGLE CYCLIZATION REACTION WHICH CONVERTS
 CC LYCOPENE TO DELTA-CAROTENE AND NEUROSPORENE TO ALPHA-ZEACAROTENE.
 CC -|- SUBCELLULAR LOCATION: Chloroplast.
 CC -|- SIMILARITY: BELONGS TO THE LYCOPENE CYCLASE FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 DR EMBL: Y14387; CAAY745.1;
 KW Oxidoreductase; NAD; Carotenoid biosynthesis; Chloroplast;
 KW Transit peptide.
 FT TRANSIT 1 ? CHLOROPLAST (POTENTIAL).
 FT CHAIN ? 526 LYCOPENE EPSILON CYCLASE.
 FT NP_BIND 108 136 NAD (POTENTIAL).
 SQ SEQUENCE 526 AA; 58886 MW; 37357C3869DBDCAF CRC64;

QY 1 MELGLVRNLISCPV-----WTFCTNLTSSK---LAYNHRVYSGSCRDYDFQVRADGG 50
 D 1 MECVGQVNGVAMAVLTRPLNRWGG--ELCOEKISFLAY--EQYESKC-----NSS 48
 QY 51 SGRSSSVAYKRGFDEDFIKAGSELFLVQMOQTQKMEKQAKLADKLPIPPFGESVMDL 110
 D 49 SGRSSCVVDREDADEEDYIKAGSQLVFVQMOQKQMDQSKLDELRLQISAGQVLDL 108
 QY 111 VVICGPAGLSLAALAAKGLKVLGPDLPFTNNYGVWDEDFKDLGLERCIEHAWKDTI 170
 D 109 VVICGPAGLALAAESAKGLKVLGPDLPFTNNYGVWDEDFKDLGLQACIEHWRTDI 168
 QY 171 VYLDNDAPVLIGRAYGRVSRHLLHEELKRCVSGSVLDSKVERITEAGDGHSLVVCEN 230
 D 169 VYLDNDPILIGRAYGRVSRHLLHEELKRCVSGSVLDSKVERITEAGDGHSLVVCEN 228
 QY 231 EIFPCRLATVASGAASGLLEYEVGGPRVCQYAYGVEVEVNNPYDPNLMVMDYRDY 290
 D 229 DWVPCRFVTVASGAASGKFLQYELGSPRVSVQYAYGVEVEVDNPPDPSLMVMDYRDY 288
 QY 291 MQQKLGQSEEBYPTFLYVMPSPRLFFETCLASKDAMPDLKRLKMSRLKTLGIQV 350
 D 289 LRHDAQSLAEKYPFLYAMPSPTRVFFETCLASKDAMPDLKRLKMSRLKTLGIQV 348
 QY 351 KYVEEESYIPVGGSLPNTQKNLAFGAASWVHPATGYSVVRSLSEAPKASYIAKILK 410
 D 349 EYEEESYIPVGGSLPNTQKNLAFGAASWVHPATGYSVVRSLSEAPKASYIAKILK 408
 QY 411 QDNSAYVVGSSAVNISMQAWSSLPKRRQRRAFFLGLGLIVQLDIEATRTFFFRFF 470
 D 409 QHYSKNMLT-SSSTPSISTQAWNTLWPKRRQRRAFFLGLGLIVQLDIEATRTFF 467
 QY 471 RLPTWMMWGLGSLSSDPLVFSYMFVLPAPNSMRSLVRHLLSDPGAVVWVAYL 527
 D 468 RVPKMMWQGLGSLSSADLMLFAFYMTIAPNDMRKGLIRHLLSDPTGATLIRTYL 524

RESULT 3

LCYB_CAPAN STANDARD; PRT; 498 AA;
 AC Q43415;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Lycopen beta cyclase, chloroplast precursor (EC 1.14.-.-).
 GN LCY1 OR CRTL.
 OS Capsicum annuum (Bell pepper).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Capsicum.
 OX NCBI_TaxID=4072;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Lamuyo; TISSUE=Fruit;
 RX MEDLINE=96045549; PubMed=7550379;
 RA Huguency P., Badillo A., Chen H.C., Klein A., Hirschberg J.,
 RA Camara B., Kuntz M.;
 RT "Metabolism of cyclic carotenoids: a model for the alteration of this
 biosynthetic pathway in Capsicum annuum chromoplasts.";
 RL Plant J. 8:417-424(1995).
 CC -|- FUNCTION: CATALYZES THE DOUBLE CYCLIZATION REACTION WHICH CONVERTS
 CC LYCOPENE TO BETA-CAROTENE AND NEUROSPORENE TO BETA-ZEACAROTENE.
 CC -|- PATHWAY: Carotenoid biosynthesis.
 CC -|- SUBCELLULAR LOCATION: Chloroplast; chromoplast.
 CC -|- SIMILARITY: BELONGS TO THE LYCOPENE CYCLASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 DR EMBL: Y14387; CAAY745.1;
 KW Oxidoreductase; NAD; Carotenoid biosynthesis; Chloroplast;
 KW Transit peptide.
 FT TRANSIT 1 ? CHLOROPLAST (POTENTIAL).
 FT CHAIN ? 526 LYCOPENE EPSILON CYCLASE.
 FT NP_BIND 108 136 NAD (POTENTIAL).
 SQ SEQUENCE 526 AA; 58886 MW; 37357C3869DBDCAF CRC64;

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: X86221; CAA60119.1; -
 InterPro: IPR000205; NAD_binding.
 Oxidoreductase; NAD; Carotenoid biosynthesis; Chloroplast;
 Transit peptide.
 TRANSIT 1 79 CHLOROPLAST (POTENTIAL).
 CHAIN 80 498 LYCOPENE BETA CYCLASE.
 NP_BIND 84 112 NAD (POTENTIAL).
 SEQUENCE 498 AA; 55610 MW; 177180CD5745F64F CRC64;

Query Match 27.7%; Score 761; DB 1; Length 498;
 Best Local Similarity 35.1%; Pred. No. 7.3e-51;
 Matches 174; Conservative 100; Mismatches 184; Indels 38; Gaps 12;

QY 43 FQVRADGGSGRS---SVAYKEGVDEEDFTKAGGSELL-FVQMOQTKSMEKQAKLADK 97
 DB 17 FGVKVSASFSSVKSRFGAKKFCGLGSRVSVCKASSALLELVPTKKNLDFELPMYD- 75
 QY 98 LPIPIFGESVMDLVVIGCGPAGLSAAEAAGLKVGLIGPD--LPFTNNYGVWEDEPKD 155
 DB 76 ---PSKGVVVDLAVVGGPAGLAVAAQVSEAGLSVCSIDPNPKLTPNPNYGVWVDEFEA 131
 QY 156 LGLERCIEHAWKDTIVYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVSGSVYLSKVER 215
 DB 132 MDLLDCLDATWSGAAYYIDDKTKDLNRPYGRVNRKQLSKMVKCILNGVHFHQAQVTK 191
 QY 216 ITEAGDGHSLVVCNEIFIPCRLATVAGSAGSKLLEYEVG-GPRVCVQATYGEVEVEN 274
 DB 192 VIHE-ESKSMCLCNDGITIATVVDATG-FSRSLVQYDKPNPG--YQVAYGILAEVEE 247
 QY 275 NPYDNLAVFYDYRD-YMQOKLOCSE--EYPTFLYVMPMSTRLEFETCLASDAMPF 331
 DB 248 HFDVKNVFWMDRWSLKNVVELKERNRIPTFLYAMPFSSNRIFLEETSLVAPRGGM 307
 QY 332 DLLKRLMSRLKTLGIQVTKVVEEWSYTPVGGSLPNTQKLNATGAASVHPATGYV 391
 DB 308 DDIOERWARLSHLGIKVKSTIEDEHCVPMGGLPVLQPVQVVGIGTAGVHPSTGYV 367
 QY 332 VRSLEAPKVASIAKILKQDNASVYVQSGSAVINISQAWNSLMPKERRORAFFLGL 451
 DB 368 ARTLAAPVPAVNAIIQYLSERS-----HSGDELSAAVWKLWPIERRRQREFFCFGM 420
 QY 452 ELIVOLDTEATRTFTFRPLTPMWWGLGSLSSFDLVFSMYMFLVAPNSRMS--- 508
 DB 421 DILLKLDLPATRRFFDAFDLEPRYWHGFLSLRFLPELVFGLSLFASHASNTSRLEIMT 480
 QY 509 -----LVRLHLS 516
 DB 481 KGTPLVHMINLLQD 496

RESULT 4
 LCYB_LYCES STANDARD; PRT; 500 AA.
 AC Q43503;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Lycopen beta cyclase, chloroplast precursor (BC 1.14.-.-).
 GN LCY1 OR CRT1-1.
 OS Lycopersicon esculentum (Tomato).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 OX NCBI_TaxID=4081;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. VF36; TISSUE=Leaf;
 RX MEDLINE=96434545; PubMed=8837512;

RA Cunningham F.X. Jr., Pogson B., Sun Z., McDonald K.A., Dellapenna D.,
 Gantt E.;
 RT "Functional analysis of the beta and epsilon lycopene cyclase enzymes
 of Arabidopsis reveals a mechanism for control of cyclic carotenoid
 formation.";
 RL Plant Cell 8:1613-1626(1996).
 CC -!- FUNCTION: CATALYZES THE DOUBLE CYCLIZATION REACTION WHICH CONVERTS
 LYCOPENE TO BETA-CAROTENE AND NEUROSPORENE TO BETA-ZEACAROTENE.
 CC -!- PATHWAY: Carotenoid biosynthesis.
 CC -!- SUBCELLULAR LOCATION: Chloroplast.
 CC -!- SIMILARITY: BELONGS TO THE LYCOPENE CYCLASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL: X86452; CAA60170.1; -
 InterPro: IPR000205; NAD_binding.
 Oxidoreductase; NAD; Carotenoid biosynthesis; Chloroplast;
 Transit peptide.
 TRANSIT 1 81 CHLOROPLAST (POTENTIAL).
 CHAIN 82 500 LYCOPENE BETA CYCLASE.
 NP_BIND 86 114 NAD (POTENTIAL).
 SEQUENCE 500 AA; 56180 MW; CF42F7D4684C04DD CRC64;

Query Match 27.3%; Score 749.5; DB 1; Length 500;
 Best Local Similarity 36.6%; Pred. No. 5.6e-50;
 Matches 169; Conservative 92; Mismatches 168; Indels 33; Gaps 11;

QY 73 GSELLFVQMOQTKSMEKQAKLADKLPPI-PFGESVMDLVVIGCGPAGLSAAEAAGL 131
 DB 52 GSSALLELVPEP-----KKNLDFELPMYDPSKGVVVDLAVVGGPAGLAVAAQVSEAGL 107
 QY 132 KYGLIGPD--LPFTNNYGVWEDEKDLGRCIEHAWKDTIVYLDNDAPVLIGRAYGRVS 189
 DB 108 SVCSIDPNKLIWPNYGVWVDEFEAMDLLDCLDATWSGAAYIDDNTAKDLHRPYGRVN 167
 QY 190 RHLLEELKRCVSGSVYLSKVERITEAGDGHSLVVCNEIFIPCRLATVAGSAGSK 249
 DB 168 RKQLSKMVKQKCTMGVFKHQAQVVIHE-ESKSMCLCNDGITIATVVDATG-FSR 225
 QY 250 LLEYEVG-GPRVCVQATYGEVEVENNPDNLMVMDYRD-YMQOKLOCSE--EYPTF 305
 DB 226 LVOYDYPYNG--YQVAYGILAEVEEHPFDVKNVFMWMDRWSHLKNTDLKERNRIPTF 283
 QY 306 LYVMPMSTRLEFETCLASDAMPFLLKRLMSRLKTLGIQVTKVVEEWSYTPVGG 365
 DB 284 LYAMPFSSNRIFLEETSLVAPRGIRIDDIQERMVARLNHLGKVKSTIEDEHCLIPMG 343
 QY 366 LPNTEOKNLAFGAASVHPATGYSVVRSLEAPKVASYAKILKQDNASVYVGGSSAV 425
 DB 344 LPVLPQVVGIGTAGVHPSTGYMVARTLAAAPVPAVNAIIQYLSERS-----HSGN 396
 QY 426 NISMQAWSLMPKERRORAFFLGLFLIVOLDIEATRTFTFRPLTPMWWGLGSL 485
 DB 397 ELSTAVKWLWPIERRRQREFFCFGMDDLKLDLPATRRFFDAFDLEPRYWHGFLSSRL 456
 QY 486 SSFDLVFSMYMFLVAPNSMR-----MSLVRLHLS 516
 DB 457 FUPELIVFGLSLFASHASNTSRFEIMTKGVTPLVNMINLLQD 498

RESULT 5
 LCYB_ARATH STANDARD; PRT; 501 AA.
 AC Q38933; Q39145;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Lycopene beta cyclase, chloroplast precursor (EC 1.14.-.-).
GN LCY1 OR LYC OR AT3G10230 OR F14P13.17.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsidae.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=96434545; PubMed=8837512;
RA Cunningham F.X. Jr., Pogson B., Sun Z., McDonald K.A., Dellapenna D.,
RA Gantt E.;
RT "Functional analysis of the beta and epsilon lycopene cyclase enzymes
RT of Arabidopsis reveals a mechanism for control of cyclic carotenoid
RT formation";
RL Plant Cell 8:1613-1626(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Wassilewskija;
RA Scolnik P.A., Bartley G.E.;
RT "Nucleotide sequence of lycopene cyclase from Arabidopsis";
RL (in) Plant Gene Register PGR95-019.
RN [3]
RP SEQUENCE FROM N.A.
RC Giuliano G., Rosati C., Santangelo G.;
RT "Gene structure and regulation of the carotenoid biosynthesis pathway
RT in Arabidopsis thaliana";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016720; PubMed=11130713;
RA Salanoubat M., Lemcke K., Rieger W., Ansoorge W., Unseld M.,
RA Fartmann B., Valle G., Bloeker H., Perez-Alonso M., Obermaier B.,
RA Delisney M., Boutry M., Grivell L.A., Mache R., Puigdomenech P.,
RA De Simone V., Choise N., Artiguenave F., Robert C., Brottier P.,
RA Wincker P., Cattolico L., Welissenbach J., Saurin W., Quetier F.,
RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,
RA Wurmbach E., Drzonek H., Erfle H., Jordan R., Bangert S.,
RA Wiedemann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G.,
RA Verzi A., D'Angelo M., Pallavicini A., Toppo S., Simonetti B.,
RA Conrad A., Hornischer K., Kauer G., Loehner T.-H., Nordstiek G.,
RA Reichelt J., Schaefer M., Schoen O., Bagues M., Terol J., Clement J.,
RA Navarro P., Collado C., Perez-Perez A., Ottenwaelder B., Duchemin D.,
RA Cooke R., Laudie M., Berger-Llauro C., Purnelle B., Masuy D.,
RA de Haan M., Maarsee A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,
RA Monfort A., Argirou A., Flores M., Liguori R., Vitale D.,
RA Mannhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W.,
RA Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,
RA Rooney T., Rizzo M., Walts A., Utterback T., Fujii C.Y., Shea T.P.,
RA Creasy T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,
RA Pai G., Miltischer J., Sellers P., Gill J.E., Feldblyum T.V.,
RA Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C.,
RA Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asamizu E.,
RA Sasamoto S., Kimura T., Idesawa K., Kawashima K., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakayama S., Nakazaki N., Shino S., Takeuchi C., Wada T.,
RA Watanabe A., Yamada M., Yasuda M., Tabata S.;
RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis
RT thaliana";
RL Nature 408:820-822(2000).
CC -1- FUNCTION: CATALYZES THE DOUBLE CYCLIZATION REACTION WHICH CONVERTS
CC LYCOPENE TO BETA-CAROTENE AND NEUROSPORINE TO BETA-ZEACAROTENE.
CC -1- SUBCELLULAR LOCATION: Chloroplast.
CC -1- SIMILARITY: BELONGS TO THE LYCOPENE CYCLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U50739; AAB53337.1; -
DR EMBL: L40176; AAA81880.1; -
DR EMBL: AF117256; AAF82388.1; -
DR EMBL: AC009400; AAF02819.1; -
DR InterPro: IPR000205; NAD_binding.
KW Oxidoreductase; NAD; Carotenoid biosynthesis; Chloroplast;
KW Transit peptide.
FT TRANSIT 1 80 CHLOROPLAST (POTENTIAL).
FT CHAIN 81 501 LYCOPENE BETA CYCLASE.
FT NP_BIND 85 113 NAD (POTENTIAL).
FT FT CONFLICT 31 31 H -> P (IN REF. 2).
FT FT CONFLICT 243 243 V -> I (IN REF. 2).
SQ SEQUENCE 501 AA; 56176 MW; C3014578D0BDC4E2 CRC64;
Query Match 27.3%; Score 748; DB 1; Length 501;
Best Local Similarity 37.9%; Pred. No. 7.3e-50;
Matches 169; Conservative 90; Mismatches 163; Indels 24; Gaps 10;
Qy 74 GSELLFVQMOTKSNKQAKLADKLPIPFGES-VMDLWIGCGPAGLSLAAEAKLGLK 132
Db 52 GSAALLDLVPET-----KKNLDFELPLYDTSKQVVDLAVGGPAGLVAQVSEAGLS 107
Qy 133 VGLI--GPDLPFTNNYGVWDEDFKDLGLERCIHAKWDTIVYLDNDAPVLIGRAYGRVSR 190
Db 108 VCSIDPSPKLWPNYGVWDEFEFAMLLDCLDTTWSGAVVYDGVKKDLRSRYGRVNR 167
Qy 191 HLLHEELLKRCVSGSVYLDKSKVERITEAGDHSVLVCENEIFPCRLATVASGAAGKL 250
Db 168 KQLSKMLQKCIITNGVKFQSHQNTNVVHE-EANSTVVCSDGVKIQASVLDATG-FSRCL 225
Qy 251 LEYEVG-GPRVCVOTAYGVEVEVNNYDNLAVFMDYRD-----YMQOKLQCESEYPT 304
Db 226 VQDKPNPG--YQVAYGVAEVDGHPFDVDMVDFMDRDKLDHLSYPELKERN--KIPT 281
Qy 305 FLVYMPSPTRLEFEETCLASKDAMPDLKRLKMLKTLGIQVTKVYEEVSYIPVG 364
Db 282 FLYAMPSSNRIEETSLVARGCLMEDIQERMAARKHLGNVKRIEEDERCVPIMG 341
Qy 365 SLNPTOKNLAFGAASAMVHPATGYSVVRSLSAPKYASVIAKILKODNSAYVVGSSA 424
Db 342 PLPLVQPVVVGIGTGMVHPSTGYMARTAAAPVANAIVRYLSPSSNLSRGLQSLA 401
Qy 425 VNLSMAWSSLSLWPKRRQRARFLGLELVQLDIEATRTFFFTPLPTWMMWGLSS 484
Db 402 -----EVRDLVPIERRRQREFFCGMDILLKLDLDATRRFFDAFFDLQPHYHGFSSR 456
Qy 485 LSSFDLVLFMSYMFVLPAPNSRMSLV 510
Db 457 LFLPELLVGLSLFASHASNTSRLEIM 482
RESULT 6
LCYB.TOBAC
ID LCYB.TOBAC STANDARD; PRT; 500 AA.
AC Q43578;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lycopene beta cyclase, chloroplast precursor (EC 1.14.-.-).
GN LCY1 OR CRT1-1.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Samsun NN; TISSUE=Leaf;
RX MEDLINE=96434545; PubMed=8837512;
RA Cunningham F.X. Jr., Pogson B., Sun Z., McDonald K.A., Dellapenna D.,

or send an email to license@sib-sib.ch.

EMBL: X98796; CAA67331.1; -

Oxidoreductase; NAD; Carotenoid biosynthesis; Chloroplast;

Transit peptide; Membrane.

TRANSIT 1 85

CHAIN 86 503 CHLOROPLAST (POTENTIAL).

NP_BIND 90 117 LYCOPENE BETA CYCLASE.

SEQUENCE 503 AA; 56910 MW; 3FD1E355EF184D98 CRC64;

Query Match 25.6%; Score 701.5; DB 1; Length 503;

Best Local Similarity 36.4%; Pred. No. 2.7e-46;

Matches 159; Conservative 87; Mismatches 160; Indels 31; Gaps 11;

QY 93 KLADKLPP- - - - - PFGEVMDLVVIGGPGAGLSLAAEAAKGLKVGIGPD 139

DB 60 QLLDLVPELKKHEFDLPYDPSKALTLDLAVVGGGLARSCTSLGG-GLSVSIDPN 118

QY 140 --LPFTNNGVWEDEPKDLGLERCIEHAWKDTIVLDNDAPVLIGRAYGRVSRHLLHEEL 197

DB 119 PKLIWPNNGVGVWEDEMDLLDCLDATWSGAIVYVDRSTKNLSRPARVNRKLNLSKM 178

QY 198 LKRCVSGVSYLDSKVERITEAGDGHSLVVCENEIFIPCLRLATVAGSAAGSKLLEYEVG- 256

DB 179 MKKCSNGVRHQATVVKAMHE-EKSYLTCSGVTIDARVLDATG-FSRCLVQYDKPY 236

QY 257 GPRVCQTAYGVEVEENNPNLMVMDYRD-YMOQKLOCSEE--EYPTFLYVMPMS 313

DB 237 NFG--YQVAYGILAEVEEHPFDVKVMDWRDLSHNGKAELNERNAKPTFLYAMPFS 294

QY 314 TRLEFEETCLASKAMPFDLLKRLKMSRLTKLGIOVTKVYEEMSWIPVGGSLPNTQKN 373

DB 295 NRIFLEETSLVARPGLKMDIQERVARLNHLGIRIKSIEEDERCVPIMGGLPVPQV 354

QY 374 LAFGAASVMVHPATGYSVVRSLSEAPKYASVIKILKQNSAVVSGQSSAVNISQWAS 433

DB 355 VGIGTAGMVHSTGTWARTAAATVANSIVQYLVSDSG---LSGN---DLSADVWK 407

QY 434 SLWPKRRQRAFFLFGLELIVQLDIEATRTFTFRPLPTWMMWGLGSSLSFDLVLF 493

DB 408 DLWPIERRQREFFCFGMDILLKLDLEGTFRFFDAFFDLEPRYWHGFLSRLPLPVP 467

QY 494 SHYMFVLAPNSMRSLV 510

DB 468 GLSLFASHASNTCKLEIM 484

RESULT 10

LCYB_SYN7

ID LCYB_SYN7 STANDARD; PRT; 411 AA.

AC Q55276;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Lycopene beta cyclase (EC 1.14.-.-).

GN CRTL OR LCY.

OS Synchococcus sp. (strain PCC 7942) (Anacystis nidulans R2).

OC Bacteria; Cyanobacteria; Chroococcales; Synchococcus.

OX NCBI_TaxID=1140;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=95003701; PubMed=7919981;

RA Cunningham F.X. Jr., Sun Z., Chamovitz D., Hirschberg J., Gantt E.;

RT "Molecular structure and enzymatic function of lycopene cyclase from

RT the cyanobacterium Synchococcus sp strain PCC7942.";

RL Plant Cell 6:1107-1121(1994).

CC -!- FUNCTION: CATALYZES THE DOUBLE CYCLIZATION REACTION WHICH CONVERTS

CC LYCOPENE TO BETA-CAROTENE AND NEUROSPORENE TO BETA-ZEACAROTENE.

CC -!- ENZYME REGULATION: INHIBITED BY THE BLEACHING HERBICIDE 2-(4-

CC METHYLPHENOXY)TRILETHYLAMINE HYDROCHLORIDE (NPTA).

CC -!- PATHWAY: Carotenoid biosynthesis.

CC -!- SIMILARITY: BELONGS TO THE LYCOPENE CYCLASE FAMILY.

CC

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

EMBL: X74599; CAA52677.1; -

InterPro: IPR00205; NAD_binding.

KW Oxidoreductase; NAD; Carotenoid biosynthesis.

FT NP_BIND 4 32 NAD (POTENTIAL).

SQ SEQUENCE 411 AA; 46085 MW; C46CC5B2E85E7AC2 CRC64;

Query Match 21.9%; Score 602; DB 1; Length 411;

Best Local Similarity 35.5%; Pred. No. 8.9e-39;

Matches 150; Conservative 61; Mismatches 165; Indels 46; Gaps 10;

QY 107 VMDLVVIGCGPAGLSLAAEAAKGLKV-GL--IGPDLPTFTNNGVWEDEPKDLGLERCIE 163

DB 1 MFDALVIGSGPAGLAAIAAELAQRGLKVQGLSPVDPFPHWENTYINGPGLDLSGLEHFG 60

QY 164 HAWKDTIVLDNDAPVLIGRAYGRVSRHLLHEELKRCVSGVSYLDSKVERITEAGDGH 223

DB 61 HRWSNCVSYF-GEAPVQHQYNYGLFDRALQHQHWRQCEQGLQWLQGLKAAAI--AHD 117

QY 224 SLVVCENEIFIPCLRLATVAGSAAGSKLLEYEVGGR- - - - - VCOTAYGVEVEV 272

DB 118 HSCV- - - - - TTAAGQELQARLVVDTTGHOAFTQPHSDAIAQAAVIGIQF 165

QY 273 ENNPYDNLVMDYR-DYMOQKLOCSEEEY--PTFLYVMPMSPTLFFETCLASKDA 328

DB 166 SQPIEPHQFVMDYRSDHL- - - - - SPEERQLPPTFLYAMDGLNDVVFVEETSLAACPA 219

QY 329 MPFDLLKRLKMSRLTKLGIOVTKVYEEMSWIPVGGSLPNTQKNLAFGAASVMVHPATG 388

DB 220 IPYDLRKQRIYQRLATRGVTVQVHQHEEYCLFPMNLPLDPLTQSVGFGGAASVMVHPAS 279

QY 389 YSVVRSLSEAPKYASVIKILKQNSAVVSGQSSAVNISQWASLSLWPKRRQRAFFL 448

DB 280 YMVGALLRRAPDLANAAGLNASSL- - - - - TTAELATQAMRGLWPTKIRKHYIQ 332

QY 449 FGLELIVQLDIEATRTFTFRPLPTWMMWGLGSSLSFDLVLFMSYMFVLAPNSMRMS 508

DB 333 FGLEKLMRFSEAQLNHHFTFFGLPKQWQYGFGLTNTLSLPELIQAMRLRFAQAPNDVRWG 392

QY 509 LV 510

DB 393 LM 394

RESULT 11

Y801_DEIRA

ID Y801_DEIRA STANDARD; PRT; 410 AA.

AC Q9RW68;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Putative carotenoid cyclase DR0801.

GN DR0801.

OS Deinococcus radiodurans.

OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;

OC Deinococaceae; Deinococcus.

OX NCBI_TaxID=1299;

RN [1]

RP SEQUENCE FROM N.A.

RX STRAIN=R1;

RX MEDLINE=20036896; PubMed=10567266;

RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,

RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,

RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,

RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,

RA Makarova K.S., Aravind L., Minton K.W., Fleischmann R.D.,

RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
 RA Fraser C.M.;
 RT "genome sequence of the radioresistant bacterium Deinococcus
 RL radiodurans R1.";
 RL Science 286:1571-1577(1999).
 CC -!- SIMILARITY: BELONGS TO THE LYCOPENE CYCLASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF001934; AAF10377.1; -
 DR TIGR: DR0801; -
 DR InterPro: IPR003042; Rng_mnoxygenase.
 DR PRINTS: PR00420; RNMNOXGNASE
 KW Hypothetical protein; Oxidoreductase; NAD; Carotenoid biosynthesis;
 KW Complete proteome.
 FT NP_BIND 11 39 NAD (POTENTIAL).
 SQ SEQUENCE 410 AA; 43161 MW; EIB3162F10F9E6AF CRC64;
 Query Match 15.7%; Score 429.5; DB 1; Length 410;
 Best Local Similarity 31.0%; Pred. No. 1.5e-25;
 Matches 135; Conservative 53; Mismatches 204; Indels 43; Gaps 13;
 QY 102 PFGESVMDLVITGCGPAGLSLAAEAKLGLKVLIG--PDLPFTNNYGVWEDEFKDLG-- 157
 DB 3 PFPASSDVLVITGGPSGTALSAAELGLDVQQLAPHPPPFATYGAW-----LGLD 56
 QY 158 ---LERCIEHAKDVIYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVE--SGVSYLDSK 212
 DB 57 PFWAGCAEQVWDVRAV-TGQPQTSLGOPYA-----LLDNAALLRLGLADMTWVEGA 110
 QY 213 VERITEAGDGHSLVVCENEIIPCLRLATVAGSAAGSKLLEYEVGSPRCVCTAYGVEVEV 272
 DB 111 ALHAERSGAGWTVYAGGERW-QTFLVVDASG--HGALVSPVPFGGAALQTAYGVAARF 167
 QY 273 ENNPYDPLMVFMDYRDYMQKLCQSEEEYPTFLVYMPSPTRLFTEETCLASKDAMPFD 332
 DB 168 RRPVPTGSMVMYDRT-PAPELKRGE---ATFLYAMHLGGDRVFEVETSILIARPAITRA 223
 QY 333 LKRRKLSRLKTLGIQVTKVVEEWSYIPVGGSLPNTPEQKLNLAFAAASWHPATGYSVV 392
 DB 224 ELRRLLARLSAQGTTPHATSEEWAFPMNAQAP-APGGVLAAYGAAAGRVHPVSGFQA 282
 QY 393 RSLSEAPKYSVIAKILKQDNSAYVVSQSSAVNISQWSSSLWPKKRKRORAFELGLE 452
 DB 283 GALSADAPGATAIATALCQKDA-----AAAGWAALWSPERRAAREVHLLGVG 330
 QY 453 LIVOLDIEATRTFTFTRLPVMMWGLASSLSLSDFLVFSMTVMFVLAPNSMRLVHRH 512
 DB 331 ALLGLERAELPHFFGTFFGLPREGQWARFLHPDPTDAGTLARTMLRVFAQTGGVRVLPARA 390
 QY 513 LLSDP--SGAVMVA 525
 DB 391 ALAQPAASGRALAAA 405
 RESULT 12
 ERIL_BRANA STANDARD; PRT; 506 AA.
 AC 065727;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Squalene monoxygenase 1.1 (EC 1.14.99.7) (Squalene epoxidase 1.1) (SE
 GN 1.1).
 DE SQP1.1.
 OS Brassica napus (Rape).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Brassica.
 OX NCBI_TaxID=3708;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Westar; TISSUE=Shoot;
 RX MEDLINE=99277589; Pubmed=10350086;
 RA Schaefer U.A.; Reed D.W., Hunter D.G., Yao K., Weninger A.M.,
 RA Tsang E.W.F., Reaney M.J.T., Mackenzie S.L., Covello P.S.;
 RT "An example of intron junctional sliding in the gene families encoding
 RT squalene monoxygenase homologues in Arabidopsis thaliana and
 RT Brassica napus.";
 RL Plant Mol. Biol. 39:721-728(1999).
 CC -!- FUNCTION: CATALYZES THE FIRST OXYGENATION STEP IN STEROL
 CC BIOSYNTHESIS AND IS SUGGESTED TO BE ONE OF THE RATE-LIMITING
 CC ENZYMES IN THIS PATHWAY.
 CC -!- CATALYTIC ACTIVITY: Squalene + AH(2) + O(2) - (S)-squalene-2,3-
 CC epoxide + A + H(2)O.
 CC -!- COFACTOR: FAD.
 CC -!- SIMILARITY: BELONGS TO THE SQUALENE MONOOXYGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AJ005931; CA006773.1; -
 DR InterPro: IPR000733; Flav_monoxygenase.
 DR InterPro: IPR000205; NAD_binding.
 DR Pfam: PF01360; Monoxygenase; 1
 KW Oxidoreductase; Flavoprotein; FAD; Transmembrane; Multigene family.
 FT TRANSMEM 3 23 POTENTIAL.
 FT TRANSMEM 47 67 POTENTIAL.
 FT TRANSMEM 447 467 POTENTIAL.
 FT NP_BIND 50 77 FAD (ADP PART) (POTENTIAL).
 SQ SEQUENCE 506 AA; 55571 MW; 32C0301F1A66CD13 CRC64;
 Query Match 5.08; Score 137.5; DB 1; Length 506;
 Best Local Similarity 24.3%; Pred. No. 0.0053;
 Matches 108; Conservative 52; Mismatches 145; Indels 139; Gaps 25;
 QY 89 EKOAKLADLPPIPFGESVMDLVVIGCGPAGLSLAAEAKLGLKVLIGDPL--PFTNNY 146
 DB 30 KKVAKLPDAATEVR-RGDADVILVAGVGSALAYALAKDGRVHVIERDMREPVR--- 85
 QY 147 GVWEDEF-----KDLGERCIE-----HAWKDTIYVL-----DNDAPV-LI 181
 DB 86 --MMGEFMPQGGRLLSKLGLEDCLGIDRQIATGLAVYKDGKALVSFFEDNDPVEPT 143
 QY 182 GRAY--GRVSRHL-----LHEELLKRCVE-----SGVSYLDSKVERTEAGDGH 223
 DB 144 GRAFYNGRFVQRLRQKASSLPTVQLEEGTVKSLEEKGVKGVTKYNS-----AGE-- 194
 QY 224 SLVVCENEIIPCLRLATVAGSAAS-----GKLLYEYGVGSPRCVCTAYGVEVE 273
 DB 195 -----ETTAFAP--LTVVCDGYSNLRSSVNDNNAEIVSYQVG-----YVSKNCQLE-- 239
 QY 274 NNYPDPLMVFMDYRDYMQKLCQSEEEYPTFLVYMPSPTRLFTEETCLASKDAM----- 329
 DB 240 -----DPEKLLKMSKPSFTMLYQIISSTDVRCVMEIPFGNIPISISNGEMAYILKNTMAPOV 295
 QY 330 PFDLLKRLMSRLKTLGIQV---TKVYEEWSYIPVGGSLPNTPEQKLNLAFAAASWHP 395
 DB 296 PPEL--RKIFLKGIDSGAOKAMPTRKMEATLS-----EKGVIVLGDAPFNRRHP 343
 QY 386 --ATGYSVV-----RSLSEAPKYSVIAKILKQDNSAYVY--SGQSSAVNIS 428
 DB 344 ATASGMVMVYLSDILRLRLLOPLRLNLSANKVSEVI-----KSFYVIRKPMSTVNTL 396
 QY 429 MQAWSSWL-----PKRRKRQAF 446

Db 280 AVADR-----LSALDVTSASIIHATHFARERWQQGFFRMLNRMLF 322
QY 455 VOLDIEATRTFTFRRLPTWMMWFLGSSLSSED 489
Db 323 LAGPADSRWRYMQRFGYGLPDLIARFYAGKLTLD 357

RESULT 15
CRTV_ERWHE STANDARD; PRT; 386 AA.
AC Q01331;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lycopene cyclase.
GN CRTV.
OS Erwinia herbicola.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Pantoea.
OX NCBI_TaxID=549;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ERO10;
RX MEDLINE=93138098; PubMed=8422926;
RA Hundle B.S., O'Brien D.A., Beyer P., Kleinig H., Hearst J.E.;
RT "In vitro expression and activity of lycopene cyclase and
beta-carotene hydroxylase from Erwinia herbicola.";
RL FEBS Lett. 315:329-334(1993).
CC -!- FUNCTION: CATALYSES THE CYCLIZATION REACTION WHICH CONVERTS
LYCOPENE TO BETA-CAROTENE.
CC -!- PATHWAY: Carotenoid biosynthesis.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; M87280; AAA64980.1; -
KW Carotenoid biosynthesis.
SQ SEQUENCE 386 AA; 43341 MW; F4A40563BFCFA980 CRC64;

Query Match 4.6%; Score 125; DB 1; Length 386;
Best Local Similarity 20.3%; Pred. No. 0.033;
Matches 86; Conservative 62; Mismatches 163; Indels 112; Gaps 18;

QY 109 DLVVIGCPA-GL---SLAAEAAKGLKVLGDLPLFTNNYGVWEDEF---KDLGLERC 161
Db 3 DLIIVGGGLANGLIANLRLQRYPOLNLLIEAGEQPGCNHTWSEFDDLTPOHAWLAPL 62
QY 162 IEHAKDTIYVLNDAPVLIGRAYGRVSRHLHEELKRCVSGSVYSDSKVERITEA-- 219
Db 63 VAHAWPGYEQFPD-----LRRRLARG--YYSITSERFAEALH 98
QY 220 ---GDGHSVLVCNEIFPCRLATV-----ASGAA--SGKLLYEYVGPRVCVQTAYG 267
Db 99 QALGE-----NWLNCVSEVLPSNVRNLANGEALLAGAVIDGRGVTVASSAMOTGYQ 149
QY 268 VEVEVE---NNPYDPLMVDYDMQKLOCSEEEYPTFLYVMPSPTRLFEEETCLA 324
Db 150 LFLGQWRLTQPHGLVPIILMDATVAQQOQYR-----FVYTLPLSADTLIEDTRYA 201
QY 325 SKDAMPDLLKRLKMLRLKTLGTVKTVYEEWSYIPV--GGSL-----PNTQKNLA 375
Db 202 NVPQRDNLALRQVTDYAHKSGWQLAQERETGCLPITLAGDIQALWADAPGVPRS--- 258
QY 376 FGAAASWHPATGYS-----VYRSISEAPKVASVIAKILKQNSAYVYVSGSSAVNISM 429
Db 259 -GMRAGLFHTTGYSLPLAVADATADSPRLGSV-----PLYQLTR 299
QY 430 QAWSSLPKE---RRQRRAFFLFLGLELIVOLDIEATRTFTFRRLPTWMMWFLGSSLS 486

Db 300 QFAERHWRQGFRRLLNMLFLAGRE-----ENRNRVMQRFYGLPEPTVERFYAGRLS 352
QY 487 SFD 489
Db 353 LFD 355

Search completed: May 21, 2003, 22:09:40
Job time : 27 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run On: May 21, 2003, 22:05:08 ; Search time 72 Seconds
(without alignments)
1513.874 Million cell updates/sec

Title: US-09-701-395A-23
Perfect score: 2743
Sequence: 1 MELLGVRNLSSCPVWTFGT.....VRHLLSPGAVMVRAYLER 529

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: SPTREMBL_21.*
- 2: sp-archaea.*
- 3: sp-bacteria.*
- 4: sp-fungi.*
- 5: sp-human.*
- 6: sp-invertebrate.*
- 7: sp-mammal.*
- 8: sp-mhc.*
- 9: sp-organelle.*
- 10: sp-phage.*
- 11: sp-plant.*
- 12: sp-rodent.*
- 13: sp-virus.*
- 14: sp-vertebrate.*
- 15: sp-unclassified.*
- 16: sp-rvirus.*
- 17: sp-bacteriap.*
- 18: sp-archeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2743	100.0	529	Q9AXK9	Q9axk9 adonis pala
2	2728	99.5	529	Q9AXL0	Q9axl0 adonis pala
3	1966.5	71.7	524	Q9LDV8	Q9ldv8 arabidopsis
4	1934	70.5	516	Q9FV43	Q9fv43 tagetes ere
5	1892	70.5	533	Q9AXK7	Q9axk7 lactuca sat
6	1892	69.0	517	Q8VWR6	Q8vwr6 spinacia ol
7	1786	65.1	437	Q8S3J4	Q8s3j4 citrus para
8	1518.5	55.4	382	Q9AXK8	Q9axk8 solanum tub
9	1113	40.6	262	Q8WLB9	Q8wlb9 citrus sine
10	783.5	28.6	504	Q9M546	Q9m546 citrus sine
11	778.5	28.4	524	Q9XGX3	Q9xgx3 citrus para
12	775.5	28.3	502	Q9AXL1	Q9axl1 adonis pala
13	755	27.5	511	Q9FV42	Q9fv42 tagetes ere
14	736	26.8	165	Q9SDP5	Q9sdp5 daucus caro
15	729	26.6	495	Q8S3C3	Q8s3c3 sandersonia
16	719.5	26.2	498	Q9FV32	Q9fv32 lycopersico

17	714.5	26.0	498	10	Q9M424	Q9m424 solanum tub
18	710.5	25.9	498	10	Q9LWA6	Q9lwa6 lycopersico
19	499	18.2	201	10	Q9M5J5	Q9m5j5 daucus caro
20	288	10.5	165	10	Q9SDS5	Q9sd5 daucus caro
21	246	9.0	165	10	Q9SDS6	Q9sd6 daucus caro
22	173	6.3	394	2	Q9KIX3	Q9kix3 bradyrhizob
23	151.5	5.5	385	2	Q93CI8	Q93ci8 xanthobacte
24	151.5	5.5	394	16	Q9K157	Q9k157 neisseria m
25	146.5	5.3	394	16	Q9JSS1	Q9jss1 neisseria m
26	145.5	5.3	386	2	Q8VUJ6	Q8vu16 pentoea agg
27	134.5	4.9	408	16	O06427	O06427 mycobacteri
28	134	4.9	393	17	Q8U4J0	Q8u4j0 pyrococcus
29	130	4.7	389	16	Q9CK99	Q9ck99 pasteurella
30	123	4.5	393	17	Q9V2B0	Q9v2b0 pyrococcus
31	123	4.5	530	10	Q9T064	Q9t064 arabidopsis
32	116.5	4.2	408	16	Q9CBA6	Q9cba6 mycobacteri
33	116.5	4.2	457	2	Q54453	Q54453 streptococc
34	115.5	4.2	299	17	Q8TV39	Q8tv39 methanopyru
35	115.5	4.2	379	16	Q8YOA5	Q8yoa5 raistonia s
36	115	4.2	704	16	Q8YRC4	Q8yrc4 anabaena sp
37	113	4.1	382	2	P94791	P94791 flavobacter
38	113	4.1	476	17	O59547	O59547 pyrococcus
39	112.5	4.1	405	16	Q9RI55	Q9ri55 streptomyce
40	112.5	4.1	1613	17	Q58907	Q58907 methanococc
41	112	4.1	413	2	Q9EXK5	Q9exk5 streptomyce
42	111.5	4.1	452	16	Q8YC38	Q8yc38 bruceella me
43	111.5	4.1	456	17	Q9HS24	Q9hs24 halobacteri
44	111.5	4.1	550	16	Q9KPA4	Q9kpa4 vibrio chol
45	110.5	4.0	381	17	O27753	O27753 methanobact

ALIGNMENTS

RESULT 1

ID	Q9AXK9	PRELIMINARY;	PRT;	529 AA.
AC	Q9AXK9;			
DT	01-JUN-2001 (TREMREL. 17, Created)			
DT	01-JUN-2001 (TREMREL. 17, Last sequence update)			
DT	01-DEC-2001 (TREMREL. 19, Last annotation update)			
DE	Lycopene epsilon-cyclase.			
OS	Adonis palaeatina.			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales;			
OC	Ranunculaceae; Adonis.			
OX	NCBI_TaxID=151078;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=IMMATURE AND DEVELOPING FLOWER BUDS;			
RX	MEDLINE=21126984; PubMed=11226339;			
RA	Cunningham F.X. Jr., Gantt E.;			
RT	"One ring or two? Determination of ring number in carotenoids by			
RT	lycopene varepsilon epsilon-cyclases";			
RL	Proc. Natl. Acad. Sci. U.S.A. 98:2905-2910(2001).			
CC	-!- COFACTOR: FAD (BY SIMILARITY).			
DR	EMBL; AF321536; AAK07432.1;			
DR	InterPro: IPR001327; FAD_pyr_redox.			
DR	PRINTS; PR00368; FADPNR.			
KW	FAD; Flavoprotein; Oxidoreductase.			
SQ	SEQUENCE 529 AA; 59170 MW; 784E4468F3D172B CRC64;			
Query Match 100.0%; Score 2743; DB 10; Length 529;				
Best Local Similarity 100.0%; Pred No. 1.3e-221;				
Matches 529; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	1	MELLGVRNLSSCPVWTFGTFRNLSSSKLAYNIHRYGSSCRVDFQVRADGGSSRSVAYK 60		
Db	1	MELLGVRNLSSCPVWTFGTFRNLSSSKLAYNIHRYGSSCRVDFQVRADGGSSRSVAYK 60		
QY	61	EGFVDEEDFTKAGSELFLVQMOTKSMKQAKLADKLPPIPFGESVMDLVWICGCPAGL 120		
Db	61	EGFVDEEDFTKAGSELFLVQMOTKSMKQAKLADKLPPIPFGESVMDLVWICGCPAGL 120		

```
QY 121 SLAAEAAKGLKVLGIPDLPTNNYGVWEDEFKDLGLERCIEHAWKDTIVYLDNDAPVL 180
|
|
|
Db 121 SLAAEAAKGLKVLGIPDLPTNNYGVWEDEFKDLGLERCIEHAWKDTIVYLDNDAPVL 180
|
|
|
QY 181 IGRAYGRVSRHLLHEELLKRCVSGSVYLDKSKVERITEAGDGHSLVVCENEIFPCRLAT 240
|
|
|
Db 181 IGRAYGRVSRHLLHEELLKRCVSGSVYLDKSKVERITEAGDGHSLVVCENEIFPCRLAT 240
|
|
|
QY 241 VASGAASGKLLLEYEVGGPRVCQTAYGVEVEENPNPDLNLMVMDYRDYMQOQLQCSEE 300
|
|
|
Db 241 VASGAASGKLLLEYEVGGPRVCQTAYGVEVEENPNPDLNLMVMDYRDYMQOQLQCSEE 300
|
|
|
QY 301 EYPTFLYVMPSPTRLFFETCLASKDAMPDCLKRLKMSRLKTLGIGQVTKVYEESYI 360
|
|
|
Db 301 EYPTFLYVMPSPTRLFFETCLASKDAMPDCLKRLKMSRLKTLGIGQVTKVYEESYI 360
|
|
|
QY 361 PVGSLPTEQKNLAFGAASWHPATGYSVVRSLSAPKYASVIAKILKQDNSAYVYSG 420
|
|
|
Db 361 PVGSLPTEQKNLAFGAASWHPATGYSVVRSLSAPKYASVIAKILKQDNSAYVYSG 420
|
|
|
QY 421 QSSAVNISQMAWSSLPKRRQRAFFLGLLEIVQLDIEATRTFFRFLPTMMWGF 480
|
|
|
Db 421 QSSAVNISQMAWSSLPKRRQRAFFLGLLEIVQLDIEATRTFFRFLPTMMWGF 480
|
|
|
QY 481 LGSSLSFDLVLFMYMFVLPAPNSMRSLVRHLLSDPSGAVMVYAYLER 529
|
|
|
Db 481 LGSSLSFDLVLFMYMFVLPAPNSMRSLVRHLLSDPSGAVMVYAYLER 529
|
|
|
RESULT 2
Q9AXL0 PRELIMINARY; PRT; 529 AA.
ID Q9AXL0
AC Q9AXL0;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Lycopene epsilon-cyclase.
OS Adonis palaeolina.
OC Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales;
OC Ranunculaceae; Adonis.
OX NCBI_TaxID=151078;
RN [1]
SEQUENCE FROM N.A.
RC TISSUE=IMMATURE AND DEVELOPING FLOWER BUDS;
RX MEDLINE=21126984; PubMed=11226339;
RA Cunningham F. X. Jr., Gantt E.;
RT "One ring or two? Determination of ring number in carotenoids by
  lycopene varepsilon epsilon-cyclases.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:2905-2910(2001).
CC -1- COFACTOR: FAD (BY SIMILARITY).
DR EMBL: AF321535; AAK07431.1;
DR InterPro: IPR001327; FAD_pyr_redox.
DR PRINTS; PR00368; FADPNR.
KW FAD; Flavoprotein; Oxidoreductase.
SQ SEQUENCE 529 AA; 59155 MW; F44597CB0A239F9B CRC64;

Query Match 99.5%; Score 2728; DB 10; Length 529;
Best Local Similarity 99.1%; Pred. No. 2.3e-220;
Matches 524; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

QY 1 MELLGVNRLISSCPVWTFGRNLSSSLAYNIHRYGSSCRVDFQVRADGGSSRSVAYK 60
|
|
|
Db 1 MELLGVNRLISSCPVWTFGRNLSSSLAYNIHRYGSSCRVDFQVRADGGSSRSVAYK 60
|
|
|
QY 61 EGFVDEEDFTKAGSELLFVQMOTQKNEKQAKLADLPPIPFGEVMDLVVIGCGPAGL 120
|
|
|
Db 61 EGFVDEEDFTKAGSELLFVQMOTQKNEKQAKLADLPPIPFGEVMDLVVIGCGPAGL 120
|
|
|
QY 121 SLAAEAAKGLKVLGIPDLPTNNYGVWEDEFKDLGLERCIEHAWKDTIVYLDNDAPVL 180
|
|
|
Db 121 SLAAEAAKGLKVLGIPDLPTNNYGVWEDEFKDLGLERCIEHAWKDTIVYLDNDAPVL 180
|
|
|
```

```
QY 181 IGRAYGRVSRHLLHEELLKRCVSGSVYLDKSKVERITEAGDGHSLVVCENEIFPCRLAT 240
|
|
|
Db 181 IGRAYGRVSRHLLHEELLKRCVSGSVYLDKSKVERITEAGDGHSLVVCENEIFPCRLAT 240
|
|
|
QY 241 VASGAASGKLLLEYEVGGPRVCQTAYGVEVEENPNPDLNLMVMDYRDYMQOQLQCSEE 300
|
|
|
Db 241 VASGAASGKLLLEYEVGGPRVCQTAYGVEVEENPNPDLNLMVMDYRDYMQOQLQCSEE 300
|
|
|
QY 301 EYPTFLYVMPSPTRLFFETCLASKDAMPDCLKRLKMSRLKTLGIGQVTKVYEESYI 360
|
|
|
Db 301 EYPTFLYVMPSPTRLFFETCLASKDAMPDCLKRLKMSRLKTLGIGQVTKVYEESYI 360
|
|
|
QY 361 PVGSLPTEQKNLAFGAASWHPATGYSVVRSLSAPKYASVIAKILKQDNSAYVYSG 420
|
|
|
Db 361 PVGSLPTEQKNLAFGAASWHPATGYSVVRSLSAPKYASVIAKILKQDNSAYVYSG 420
|
|
|
QY 421 QSSAVNISQMAWSSLPKRRQRAFFLGLLEIVQLDIEATRTFFRFLPTMMWGF 480
|
|
|
Db 421 QSSAVNISQMAWSSLPKRRQRAFFLGLLEIVQLDIEATRTFFRFLPTMMWGF 480
|
|
|
QY 481 LGSSLSFDLVLFMYMFVLPAPNSMRSLVRHLLSDPSGAVMVYAYLER 529
|
|
|
Db 481 LGSSLSFDLVLFMYMFVLPAPNSMRSLVRHLLSDPSGAVMVYAYLER 529
|
|
|
RESULT 3
Q9LDV8 PRELIMINARY; PRT; 524 AA.
ID Q9LDV8
AC Q9LDV8;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Lycopene epsilon cyclase (Putative lycopene epsilon cyclase).
GN AT5G57030.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=20181125; PubMed=10718197;
RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Kotani H.,
  Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence
  features of the regions of 3,076,755 bp covered by sixty P1 and TAC
  clones.";
RL DNA Res. 7:31-63(2000).
RN [2]
SEQUENCE FROM N.A.
RA Giuliano G., Rosati C., Santangelo G.;
RT "Gene structure and regulation of the carotenoid biosynthesis pathway
  in Arabidopsis thaliana.";
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
  Dale J.M., Gibson H.A., Goldsmith A.D., Jiang P.X., Lee J.M.,
  Quach H.L., Tang C., Toriumi M., Yu G., Bowser L., Carninci P.,
  Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T., Kamiya A.,
  Karlin-Neumann G., Kawai J., Kim C., Koesema E., Lam B., Lin J.,
  Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
  Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Tracy S.E.,
  Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT "Full length cDNA of gene MHM17.16/AT5G57030 (GI:8777443).";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Etgu P., Lee J.M.,
  Toriumi M., Yu G., Brooks S., Chao Q., Chen H., Karlin-Neumann G.,
  Kim C., Lam B., Miranda M., Nguyen M., Palm C.J., Shinn P.,
```


DR PRINTS; PR00368; FADPNR.
KW FAD; Flavoprotein; Oxidoreductase.
SQ SEQUENCE 533 AA; 59832 MW; 7A5806BE758AE7B2 CRC64;

Query Match 70.5%; Score 1934; DB 10; Length 533;
Best Local Similarity 68.3%; Pred. No. 1.1e-153;
Matches 375; Conservative 57; Mismatches 77; Indels 40; Gaps 5;

QY 1 MELLGVRLN-----ISSCPVMT-----FGTRNLSSSLAYNIHRYGSSCR 40
DB 1 MECFGARNMATMAVFCPRFTDCNIRKPSLLKQRFNLSASSSLRQIKCSAKSDRCV 60

QY 41 VDFOVADGGSSRSVAYKEGF--VDEEDFIKAGSELFLVQMOQTKSMEKQAKLADKL 98
DB 61 VD-----KQGISVADEEDYVYKAGSELFLVQMOQTKSMEKQAKLADKL 103

QY 99 PPIPFESVMDLVVIGGPGAGLSLAAEAALGLKVLGIGLDPDPTNNYGVWEDEFKDLGL 158
DB 104 AQPIGNCILDLVIGGPGAGLALAAESAKLGLNVLGIGLDPDPTNNYGVWQDEFGLGL 163

QY 159 ERCIEHAWKDTIVYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVSGVSYLDSKVERITE 218
DB 164 EGCIEHAWKDTIVYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVSGVSYLDSKVERITE 223

QY 219 AGDGHSLWCENEIFIPCRLATVAGSAGKLLLEYEVGGPRVCVQTAIGVEVEYENPNYD 278
DB 224 APNGYSLIECEGNITIPCRLATVAGSAGKLLLEYEVGGPRVCVQTAIGVEVEYENPNYD 283

QY 279 PNLVMDYRDYMOOKLQCESEETFLYVMPMSPTLRFEEETCLASKDAMPFLLKRL 338
DB 284 PDLVMDYRDYMOOKLQCESEETFLYVMPMSPTLRFEEETCLASKDAMPFLLKRL 343

QY 339 MSRLKTLGIVQTKVYEEWSYIPVGGSLPNTQKNLAFAGAAASVHPATGYSVVRSLSSEA 398
DB 344 MSRLKANGIRITRYEWSYIPVGGSLPNTQKNLAFAGAAASVHPATGYSVVRSLSSEA 403

QY 399 PKYASVIAKILKQDNSAYVYVGGSSAVNISQWSSLPWPKRQRQRAFFLGLVOLD 458
DB 404 PNYAAVIAKILKQDNSAYVYVGGSSAVNISQWSSLPWPKRQRQRAFFLGLVOLD 462

QY 459 IEARTFRFRFRFLPTMMWGLGSSLSFDLVLFMSYMFVLPAPNSMRMSLVRLHSDPS 518
DB 463 LEGTRFRFRFRFLPTMMWGLGSSLSFDLVLFMSYMFVLPAPNSMRMSLVRLHSDPS 522

QY 519 GAVMVRAYL 527
DB 523 GATWVAYL 531

RESULT 6
Q8VWR6 ID Q8VWR6 PRELIMINARY; PRT; 517 AA.
AC Q8VWR6
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Lycopene epsilon-cyclase.
GN LEC.
OS Spinacia oleracea (Spinach).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllales; Caryophyllales; Chenopodiaceae; Spinacia.
OX NCBI_TaxID=3562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. MELODY; TISSUE=LEAF;
RA Desouza M.L., Kollmann S.R., Schroeder W.A.;
RT "Production of Lutein in Microorganisms";
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF463497; AAL69394.1; -
SQ SEQUENCE 517 AA; 57911 MW; 0B0EABB09376F8A4 CRC64;

Query Match 69.0%; Score 1892; DB-10; Length 517;

Best Local Similarity 71.4%; Pred. No. 3.6e-150;
Matches 355; Conservative 61; Mismatches 75; Indels 6; Gaps 1;

QY 37 SSCRVDFOVADGGSSRS-----VAYKEGFVDEEDFIKAGSELFLVQMOQTKSMEK 90
DB 19 SNVVESSRRSGSVLSANSSSCVIAPEDFANEEDFIKAGSELFLVQMOQKAMDC 78

QY 91 QAKLADKLPIPFESVMDLVVIGGPGAGLSLAAEAALGLKVLGIGLDPDPTNNYGVWE 150
DB 79 YSKISDKLRQISDANELLDMVIGGPGAGLALAAESAKLGLKVLGIGLDPDPTNNYGVWE 138

QY 151 DEFKDLGLERCIEHAWKDTIVYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVSGVSYLD 210
DB 139 DEFALGLGCGCIEHAWKDTIVYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVSGVSYLD 198

QY 211 SKVERITEAGDGHSLWCENEIFIPCRLATVAGSAGKLLLEYEVGGPRVCVQTAIGVEV 270
DB 199 AKVENIMEGPDGHRVLACERGVTPCRLVTVAGSAGKLLLEYEVGGPRVCVQTAIGVEV 258

QY 271 EVENNPYDNLVMDYRDYMOOKLQCESEETFLYVMPMSPTLRFEEETCLASKDAMP 330
DB 259 EVENSYPDNLVMDYRDYMOOKLQCESEETFLYVMPMSPTLRFEEETCLASKDAMP 318

QY 331 FDLKRLKMSRLKTLGIVQTKVYEEWSYIPVGGSLPNTQKNLAFAGAAASVHPATGY 390
DB 319 FDLKRLKMSRLKTLGIVQTKVYEEWSYIPVGGSLPNTQKNLAFAGAAASVHPATGY 378

QY 391 VVRSLSAPKVASVIAKILKQDNSAYVYVGGSSAVNISQWSSLPWPKRQRQRAFFLFG 450
DB 379 VVRSLSAPKVASVIAKILKQDNSAYVYVGGSSAVNISQWSSLPWPKRQRQRAFFLFG 438

QY 451 LELIVQDLIEATRTFRFRFLPTMMWGLGSSLSFDLVLFMSYMFVLPAPNSMRMSLV 510
DB 439 LSLIVQDLIEATRTFRFRFLPTMMWGLGSSLSFDLVLFMSYMFVLPAPNSMRMSLV 498

QY 511 RHLLSDPSGAVMVRAYL 527
DB 499 RHLLSDPSGAVMVRAYL 515

RESULT 7
Q8S3J4 ID Q8S3J4 PRELIMINARY; PRT; 437 AA.
AC Q8S3J4
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Lycopene epsilon-cyclase.
OS Citrus paradisi (Grapefruit).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Sapindales; Rutaceae; Citrus.
OX NCBI_TaxID=37656;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. DUNCAN;
RA Costa M.C., Otoni W.C., Moore G.A.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF486650; AAL92114.1; -
SQ SEQUENCE 437 AA; 49012 MW; 5BECC880EF76EEC4 CRC64;

Query Match 65.1%; Score 1786; DB 10; Length 437;
Best Local Similarity 77.4%; Pred. No. 2.2e-141;
Matches 333; Conservative 44; Mismatches 53; Indels 0; Gaps 0;

QY 98 LPPIPFESVMDLVVIGGPGAGLSLAAEAALGLKVLGIGLDPDPTNNYGVWEDEFKDLG 157
DB 6 LPPISIGNGILDLVIGGPGAGLALAAESAKLGLNVLGIGLDPDPTNNYGVWEDEFKDLG 65

QY 158 LERCIEHAWKDTIVYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVSGVSYLDSKVERIT 217
DB 66 LEGCIEHAWKDTIVYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVSGVSYLDSKVERIT 125

[illegible]

[1]

RP SEQUENCE FROM N.A.
RA Xu C.J., Zhang S.L.;
RT "Molecular cloning of lycopene beta-cyclase gene from orange (Citrus
RL sinensis).";
DR EMBL; AF240787; AAF44700.2; -
DR InterPro; IPR00205; NAD_binding.
SQ SEQUENCE 504 AA; 56460 MW; 7C12B2BFB646481C CRC64;

Query Match 28.6%; Score 783.5; DB 10; Length 504;
Best Local Similarity 38.3%; Pred. No. 4.1e-57;
Matches 172; Conservative 94; Mismatches 160; Indels 23; Gaps 10;

Qy 69 FIKAGGSELL-FVQMOTKSMKQAKLADKLPPFPFGESVMDLVVIGCGPAGLSAAEAA 127
Db 53 FIKASSALLBELVPETKKELEFELPMYD-----PSKGLVVDLAVVGGGPAGLAAQVVS 107
Qy 128 KLGLKVLGI--GPDLPFTNNYGVWEDEFKOLGLERCIEHAWKDTIVYLDNDAPVLIGRAY 185
Db 108 EAGLSVCSIDPSPKLIWPNNTYGVWVDEFEAMDLLDCLDTTWSGAVVHDDDTKKDLDRPY 167
Qy 186 GRVSRHLLHEELLKRCVSGSYLDSKVERITEAGDGHSLVVCENEIFIPCLATVASGA 245
Db 168 GRVNRKLLKSMQKQITNGVKFHOAKVIKVIHE-ESKSLICNDGVTIQAQVLDATG- 225
Qy 246 ASGKLLYEVG-GRVVCVQTAYGEVEVENNPDPNLMVFMDYRD---YMOQKLCQSEEE 301
Db 226 FSRCLVQYDKPYNPG--YQVAYGILAEVEHPFDLDMKVMFMDWRDLSHNNSELKEANSK 283
Qy 302 YPTFLYVMPSPTRFLFEETCLASKDAMPFDLLKRLKMSRLTKTGIOVTKYVEEWSYIP 361
Db 284 IPTFLYAMPFSNRIFFLEETSLVAPRGVPMKDIQERMVARLKLHLGKIVRSTEEDEHCVIP 343
Qy 362 VGGSLPNTQKNLAFGAASVMHPATGYSVVRSLSSEAPKYASVIAKILKQDNSAYVWSGQ 421
Db 344 MGGPLVPLPQVRVGGTAGMVHPSTGVMWARTLAAPIVANAIVRSLSRSRDS---ISGH 400
Qy 422 SSAYNISQAWSSLLWPKERKQRAFFLGLLEIVOLDIEATFTFTFRPLRTMWMWGL 481
Db 401 ----KLSAEVWKDLWPIERRRQREFFCFGMDILLKLDLPATRRFFDAFDLEPRYWHGFL 456
Qy 482 GSSLSFDLVLFMSYMFVLAIPNSMRSLV 510
Db 457 SSRFLPELVGLSLFSSHASNTSRLEIM 485

RESULT 11
Q9XGX3 PRELIMINARY; PRT; 524 AA.
AC Q9XGX3;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Lycopene cyclase.
OS Citrus paradisi (Grapefruit).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Sapindales; Rutaceae; Citrus.
OX NCBI_TaxID=37656;
RN [1]
RP SEQUENCE FROM N.A.
RA Costa M.C., Moreira C.D., Melton J.R., Otoni W.C., Moore G.A.;
RT "Developmental expression of carotenoid genes in Citrus.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF152246; AAD38049.2; -
DR InterPro; IPR00205; NAD_binding.
SQ SEQUENCE 524 AA; 58523CF5B3B9208C CRC64;

Query Match 28.4%; Score 778.5; DB 10; Length 524;
Best Local Similarity 38.3%; Pred. No. 1.1e-56;
Matches 172; Conservative 93; Mismatches 161; Indels 23; Gaps 11;

Qy 69 FIKAGGSELL-FVQMOTKSMKQAKLADKLPPFPFGESVMDLVVIGCGPAGLSAAEAA 127

Db 73 FIKASSALLBELVPETKKELEFELPMYD-----PSKGLVVDLAVVGGGPAGLAAQVVS 127
Qy 128 KLGLKVLGI--GPDLPFTNNYGVWEDEFKOLGLERCIEHAWKDTIVYLDNDAPVLIGRAY 185
Db 128 GAGLSVCSIDPSPKLIWPNNTYGVWVDEFEAMDLLDCLDTTWSGAVVHDDDTKKDLNRPY 187
Qy 186 GRVSRHLLHEELLKRCVSGSYLDSKVERITEAGDGHSLVVCENEIFIPCLATVASGA 245
Db 188 GRVNRKLLKSMQKQITNGVKFHOAKVIKVIHE-ESKSLICNDGVTIQAQVLDATG- 245
Qy 246 ASGKLLYEVG-GRVVCVQTAYGEVEVENNPDPNLMVFMDYRD---YMOQKLCQSE--EE 301
Db 246 FSRCLVQYDKPYNPG--YQVAYGILAEVEHPFDLDMKVMFMDWRDLSHNNSELKEANSK 303
Qy 302 YPTFLYVMPSPTRFLFEETCLASKDAMPFDLLKRLKMSRLTKTGIOVTKYVEEWSYIP 361
Db 304 IPTFLYAMPFSNRIFFLEETSLVAPRGVPMKDIQERMVARLKLHLGKIVKASIEDEHCVIP 363
Qy 362 VGGSLPNTQKNLAFGAASVMHPATGYSVVRSLSSEAPKYASVIAKILKQDNSAYVWSGQ 421
Db 364 MGGPLVPLPQVRVGGTAGMVHPSTGVMWARTLAAPIVANAIVRSLSRSRDS---ISGH 420
Qy 422 SSAYNISQAWSSLLWPKERKQRAFFLGLLEIVOLDIEATFTFTFRPLRTMWMWGL 481
Db 421 ----KLSAEVWKDLWPIERRRQREFFCFGMDILLKLDLPATRRFFDAFDLEPRYWHGFL 476
Qy 482 GSSLSFDLVLFMSYMFVLAIPNSMRSLV 510
Db 477 SSRFLPELVGLSLFSSHASNTSRLEIM 505

RESULT 12
Q9AXL1 PRELIMINARY; PRT; 502 AA.
AC Q9AXL1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Lycopene beta-cyclase.
OS Adonis palaeatina.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; Ranunculales;
OC Ranunculaceae; Adonis.
OX NCBI_TaxID=151078;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=21126984; PubMed=11226339;
RA Cunningham F.X. Jr., Gantt E.;
RT "One ring or two? Determination of ring number in carotenoids by
RT lycopene varepsilon epsilon-cyclases.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:2905-2910(2001).
DR EMBL; AF321534; AAK07430.1; -
DR InterPro; IPR00205; NAD_binding.
SQ SEQUENCE 502 AA; 56575 MW; B973FD43794D3F1B CRC64;

Query Match 28.3%; Score 775.5; DB 10; Length 502;
Best Local Similarity 36.3%; Pred. No. 1.9e-56;
Matches 181; Conservative 99; Mismatches 166; Indels 53; Gaps 14;

Qy 23 LSSSKLAYNIHRYGS----SCRVDQVRADGGSGSRSSVAYKEGFVDEEDFIKAGSELL 78
Db 27 VSTSKLQNVFRIASRNTHPCR-----NGT-----VKRGSALL 60
Qy 79 FVQMOTKSMKQAKLADKLPPFPFGESVMDLVVIGCGPAGLSAAEAAKILKVLGI- 136
Db 61 ----ELVPETKKELEFDPAYDPSRGIVVDLAVVGGGPAGLAAQVSEAGLIVCSID 115
Qy 137 -GPDLPFTNNYGVWEDEFKOLGLERCIEHAWKDTIVYLDNDAPVLIGRAYGRVSRHLLHE 195
Db 116 FSPKLIWPNNTYGVWVDEFEAMDLLDCLDTTWSGAVVYTDNNSKKYLDRIYGRVNRKQLKS 175
Qy 196 ELLKRCVSGSYLDSKVERITEAGDGHSLVVCENEIFIPCLATVASGAASGKLLEYEV 255

```
Db 176 KMLQKVTNGVHFHQAQKIVKVIHE-ESKSLICNDGITTINATVVLDTG-FSRCLVQYDK 233
QY 256 G-GPRVCVCTAGYGEVEENNNPDNLMVMDYRD-YMOOKLQCSSE--EYTELVMVM 311
Db 234 PYNPG--YQVAGINAEVEHEFDLKDLMFMDWRSHLNKELELDKNRKIPTFLYAMFF 291
QY 312 STRLFFETCTCLASKADAMPFDLLKRLKMSRLKTLGIQVTKVYEEWYSYIPVGGSLPNTBQ 371
Db 292 SSTKIFLETSVLRPGLFQIERMAYRLKHLGIKVKISIEDERCIVPMGGPLPVLPO 351
QY 372 KNLAGAASVHPATGYSVRSLSLSEAPKYASVIAKILKODNSAYVVSQSSAVNISMQA 431
Db 352 RVVGIGTAGMHPSTGYMARTLAAPVVAKISVQYIGSDRS---LSGN----ELSAEV 404
QY 432 WSLMPKRRQRAFFLFGLELIVQIDIEATRTFFRTFLPTMMWGLSSLSFSLV 491
Db 405 WKDLPIERRRQREFFCFGMDILLKDLQGTTRFFDAFFDLPHYWHGFLSKRLFPELL 464
QY 492 LFSMTMFLVAPNSMRSLV 510
Db 465 FFLGLSLFASHNASRIEIM 483
```

```
RESULT 13
Q9FV42
ID Q9FV42 PRELIMINARY; PRT; 511 AA.
AC Q9FV42;
DT 01-MAR-2001 (TremBLrel. 16, Created)
DT 01-MAR-2001 (TremBLrel. 16, Last sequence update)
DE Beta cyclase.
OS Tagetes erecta (African marigold).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids II; Asterales; Asteraceae; Asteroidae;
OC Heleniaceae; Tagetes.
OX NCBI_TaxID=13708;
RN [1]
RP SEQUENCE FROM N.A.
RA Moehs C.P., Tian L., DellaPenna D.;
RT "Analysis of carotenoid biosynthetic gene expression during marigold
RT petal development.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -I- COFACTOR: FAD (BY SIMILARITY).
DR EMBL; AF251017; AAG10429.1;
DR InterPro; IPR001327; FAD_pyr_redox.
DR InterPro; IPR000205; NAD_binding.
DR InterPro; IPR001100; Pyr_redox.
DR PRINTS; PR00368; FADPNR.
DR PRINTS; PR00411; PNDRDTASEI.
KW FAD; Flavoprotein; Oxidoreductase.
SQ SEQUENCE 511 AA; 57831 MW; C72CEF6C3F2EF163 CRC64;
```

```
Query Match 27.5%; Score 755; DB 10; Length 511;
Best Local Similarity 36.5%; Pred. No. 1e-54;
Matches 181; Conservative 97; Mismatches 188; Indels 30; Gaps 15;

QY 25 SKLAYNIHRYG--SSCRVDFOVRADGGSGSSSVAYKEGFVDEEDFIKAGSELL-FVQ 81
Db 17 SNKFNAGNLNQLNQSKSQFO---DFRGPKKS-QFKLG---QKYCVKASSALLELP 69
QY 82 MOOTKSMKQAKLADLPPIPFGEVMDLVVIGCGPAGLSLAAEAAKGLKVGLI--GPD 139
Db 70 EIKKENLDFDLPYD-----PSRNVVDLVVVGSGSLAVAQVQSEAGLTVCSDIDPSK 124
QY 140 LPFTNYGWEDEFKDLGLERCIEHAWKDTIVYLDNDAPVLIGRAYGRVSRHLLHEELK 199
Db 125 LIWPNNGYGVWDEFAEMDLLDLDLTWSSAVYIDEKSTKSLNRPYARNRQLKTKMLQ 184
QY 200 RCVSGSVLDSKVRITEAGDGHSLVVCENEIFIPCLRLATVASGAASKLLEYEGV-GP 258
Db 195 KCIANGVRFHQAQKIVKVIHE-ELKSLICNDGVTIQTATLVLDATG-FSRSLVQYDKPYNP 242
```

```
QY 259 RVCVQYAGYGEVEENNNPDNLMVMDYRD-YMOOKLQ--CSEEBVPTFLYVMPKSPTR 315
Db 243 G--YQVAYGILAEVEHEPFDVDMKLMFMDWRDSDLQNLKARNRIPTFLYAMPFSSTR 300
QY 316 LFEETCTCLASKADAMPFDLLKRLKMSRLKTLGIQVTKVYEEWYSYIPVGGSLPNTBQ 375
Db 301 IFLEETSLVARPGLKMDIQRMAYRLKHLGIKVKISIEDERCIVPMGGPLPVLPO 360
QY 376 FGAASVHPATGYSVRSLSLSEAPKYASVIAKILKODNSAYV-VSGSSAVNISMQAWS 434
Db 361 IGTAGMHPSTGYMARTLAAPVVAKISVQYIGSDRS---LSGN----ELSAEV 404
QY 435 LMPKRRQRAFFLFGLELIVQIDIEATRTFFRTFLPTMMWGLSSLSFSLV 494
Db 417 LMPIERRRQREFFCFGMDILLKDLQGTTRFFDAFFDLPHYWHGFLSKRLFPELL 476
QY 495 MYMFLVAPNSMRSLV 510
Db 477 LSLFGHASNTCRVEIM 492
```

```
RESULT 14
Q9SDP5
ID Q9SDP5 PRELIMINARY; PRT; 165 AA.
AC Q9SDP5;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TremBLrel. 17, Last annotation update)
DE Lycopene epsilon-cyclase (Fragment).
GN LYC-E.
OS Daucus carota (Carrot).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids II; Apiales; Apiaceae; Daucus.
OX NCBI_TaxID=4039;
RN [1]
RP SEQUENCE FROM N.A.
RA Chen D., Chen H., Wang Y.;
RT "Daucus carota lycopene epsilon-cyclase.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF212130; AAF23013.1;
FT NON_TER 1
FT NON_TER 165
SQ SEQUENCE 165 AA; 18455 MW; F61CDAB0B4B53861 CRC64;

Query Match 26.8%; Score 736; DB 10; Length 165;
Best Local Similarity 83.6%; Pred. No. 7.1e-54;
Matches 138; Conservative 9; Mismatches 18; Indels 0; Gaps 0;
```

```
QY 146 YGVWDEDFKDLGLERCIEHAWKDTIVYLDNDAPVLIGRAYGRVSRHLLHEELKRCV 205
Db 1 YGVWDEFDLGLGEGIEHWRDTIVYLDGDPIMIGRAYGRVSRHLLHEELKRCV 60
QY 206 VSYLDSKVRITEAGDGHSLVVCENEIFIPCLRLATVASGAASKLLEYEGVGRVCVQTA 265
Db 61 VSYLSSKVEKITEAGDGHSLVVCENINIVIPCLRLATVASGAASKLLEYEGVGRVSVQTA 120
QY 266 YGVEVEENNNPDNLMVMDYRDYMQOKLQCSSEYPTFLYVMP 310
Db 121 YGVEVEENNNPDNLMVMDYRDYTKQVPGMEAYEPTFLYAMP 165
```

```
RESULT 15
Q8S3C3
ID Q8S3C3 PRELIMINARY; PRT; 495 AA.
AC Q8S3C3;
DT 01-JUN-2002 (TremBLrel. 21, Created)
DT 01-JUN-2002 (TremBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE Beta-lycopene cyclase.
OS Sanderionia aurantiaca (Christmas-bells) (Chinese-lantern lily).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```

OC Spermatophyta; Magnoliophyta; Liliopsida; Liliales; Liliaceae;

OC Sandersonia.

OX NCBI_TaxID=61864;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE-PETAL;

RA Nielsen K.M.;

RT "Carotenoids in Sandersonia.";

RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF489520; AAL92175.1;

SQ SEQUENCE 495 AA; 55692 MW; DAD3FE7C5FB872C1 CRC64;

Query Match

26.68; Score 729; DB 10; Length 495;

Best Local Similarity 35.78; Pred. No. 1.5e-52;

Matches 166; Conservative 92; Mismatches 173; Indels 34; Gaps 11;

```
QY 70 IKAGGSELL-FVQMOTKSMKQAKLADKLPPIPFGESVMDLVVIGGPGAGLSLAAEAAK 128
Db :| | | | | : : : : | | | | | :| | | | | :| | | | | :| |
QY 45 VVAKSSALLELVPEVKKENLDMELPLYD-----PSKSLTVDLAVVGGPGAGLAVAQVVSQ 99
Db :| | | | | :| | | | | :| | | | | :| | | | | :| |
QY 129 LGLKVGLI--GPDLPFTNNTGVWDEPKDLGLERCIEHAWKDTIIVYLDNDAPVLIGRAYG 186
Db :| | | | | :| | | | | :| | | | | :| | | | | :| |
QY 100 SGLSVCSIDPSKLIWPNNTGVWVDEFEAMDLLDCLDASWPGAVVYLDSTKLLDRPYA 159
Db :| | | | | :| | | | | :| | | | | :| | | | | :| |
QY 187 RVSRHLLHEELLKRCVESGYSLDSKYVERITEAGDGHSLVVCENEIIPICRLATVASGAA 246
Db :| | | | | :| | | | | :| | | | | :| | | | | :| |
QY 160 RVNRKQLKSMHMKCVANGVFHQAVKVIHE-EAKSNLICNDGVTIQARVVLDTAG-F 217
Db :| | | | | :| | | | | :| | | | | :| | | | | :| |
QY 247 SGKLEYEVG-GPRVCYQTAYGVEVEYENNPYPNLMVFMDYRDYM---OOKLQCSSEY 302
Db :| | | | | :| | | | | :| | | | | :| | | | | :| |
QY 218 SRCIVQYDKKPNPG--YQVAYGILAQVEEHPFDLDKAVFMDWRDMLRDGDKMDNRRI 275
Db :| | | | | :| | | | | :| | | | | :| | | | | :| |
QY 303 PTELYVMPMSPTRLFFETCLASKDAMPFDLLKRLKMSRLKTLGIQVTKYVEEWSYIPY 362
Db :| | | | | :| | | | | :| | | | | :| | | | | :| |
QY 276 PTELYAMPFSSERIFLEETSILVARPGLAMEDIQERWVARLHLGIRVKSTXXDERCIIIP 335
Db :| | | | | :| | | | | :| | | | | :| | | | | :| |
QY 363 GGSLPNTEQKNLAFGAASMHVPATGYVVVRSLSSEAPKYASVIATILKQDNSAYVVSQS 422
Db :| | | | | :| | | | | :| | | | | :| | | | | :| |
QY 336 GGPLPVLQVRVIGGTAGVHFSTGYWVARTIAAPIVAGSIVRYLSSNRG---ISGD- 391
Db :| | | | | :| | | | | :| | | | | :| | | | | :| |
QY 423 SAVNISMQAWSLWPKRKQRAFFLGLLEIVQLDIEATRTFFTRFRLPTWMMWGLG 482
Db :| | | | | :| | | | | :| | | | | :| | | | | :| |
QY 392 ---GISARVWKDLWPITERRRQREFFCFGMDILLKLDLQGTFRFFDAFFDLEPHYWHGFLS 448
Db :| | | | | :| | | | | :| | | | | :| | | | | :| |
QY 483 SSLSSFDLVLFMSMFIAPNSMRMS-----LVRHLLSD 516
Db :| | | | | :| | | | | :| | | | | :| | | | | :| |
QY 449 SRLFLPELVFGLSLFGHASNTORLEIMAKGSLPLVHMVNNLLQD 493
Db :| | | | | :| | | | | :| | | | | :| | | | | :| |
```

Search completed: May 21, 2003, 22:11:08

Job time : 78 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 21, 2003, 22:17:23 ; Search time 2143 Seconds
(without alignments)
3997.859 Million cell updates/sec

Title: US-09-701-395A-23
Perfect score: 2743
Sequence: 1 MELLGVRLNLISSCPVTFGT.....VRHLLSDPSGAVMVRAYLER 529

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2.1/USPOT_spool/US09701395/runat_15052003_165751_3370/app_query.fasta_1.711
-DB=EST -QFMT=fastap -SURFIX=first -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09701395 -CGN_1_1306.@runat_15052003_165751_3370 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=30 -DELOP=0.5 -DELEXT=7

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_rrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_othr:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1048.5	38.2	755	14	BQ856148 QGB28L15.
2	899	32.8	664	14	BQ623342 USDA-PP-0
3	868.5	31.7	658	12	BG592655 EST491333
4	860	31.4	661	14	BQ994803 QGF8D16.Y
5	843.5	30.8	719	10	AV931085 AV931085
6	793.5	28.9	665	10	AV926200 AV926200
7	728.5	26.6	610	14	BQ115144 EST601720
8	701	25.6	541	12	BG241188 OVI_39.CO
9	695	25.3	548	10	AW906884 EST343007
10	619.5	22.6	571	10	BE231430 SsS0941.S
11	602.5	22.0	614	10	AV921185 AV921185
12	600	21.9	557	10	AV521661 AV521661
13	599	21.8	465	10	AW821224 SsS0040.S
14	598.5	21.8	668	10	AV822597 AV822597
15	581.5	21.2	617	13	BJ482945 BJ482945
16	579.5	21.1	512	10	AV916064 AV916064
17	575	21.0	652	13	BI096991 SCUM33-DS
18	545	19.9	871	10	BE456098 HVSNEG001
19	516	18.8	655	12	BG240869 OVI_39.CO
20	506.5	18.5	841	17	BH446159 BOGJF81TR
21	498	18.2	442	10	AV917261 AV917261
22	498	18.2	673	13	BM449133 DSA032D12
23	460	16.8	341	10	BE318207 NF036F05L
24	459.5	16.8	622	12	BF636131 NF079H03D
25	451	16.4	325	12	BF067883 ST84f04.Y
26	425.5	15.7	420	12	BF145063 SsS0752.S
27	420	15.3	599	13	BI992439 1020062B0
28	406.5	14.8	430	10	AW599513 ga87e04.Y
29	394	14.4	547	12	BG126518 EST472164
30	390	14.2	692	12	BG127631 EST473193
31	384	14.0	590	9	AI485631 EST243952
32	375	13.7	443	14	BQ872792 QG116K13.
33	374.5	13.7	743	17	BH598608 BOGVE83TF
34	368	13.4	381	14	BU012051 QGJ17010.
35	354.5	12.9	573	12	BF423984 sso2q08.Y
36	335	12.2	649	13	BI921781 EST541684
37	334	12.2	854	14	BQ509489 EST616904
38	333.5	12.2	552	12	BG524479 42-95.Ste
39	330	12.0	487	17	BH233819 1006175H1
40	327	11.9	611	13	BI422789 EST533455
41	326.5	11.9	336	12	BF145076 SsS0766.S
42	322	11.7	404	10	AV803737 AV803737
43	319.5	11.6	507	14	BQ872037 QG113H23.
44	319.5	11.6	526	14	BQ980868 QGEL2D21.
45	318	11.6	601	12	BG646437 EST508056

ALIGNMENTS

RESULT 1
BQ856148
LOCUS
DEFINITION QGB28L15.Yg.abl QG.ABCDI lettuce salinas Lactuca sativa cDNA clone
ACCESSION BQ856148
VERSION BQ856148.1 GI:22241613
KEYWORDS EST.
SOURCE Lactuca sativa.
ORGANISM Lactuca sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids II; Asterales; Asteraceae; Lactuceae;
Lactuca.
REFERENCE 1 (bases 1 to 755)
755 bp mRNA linear EST 14-AUG-2002

AUTHORS

Kozik, A., Micheltore, R.W., Knapp, S., Matvienko, M., Rieseberg, L., Lin, H., van Damme, M., Lavelle, D., Chevalier, P., Ziegler, J., Ellison, P., Kolkman, J., Siabaugh, M.S., Livingston, K., Zhou, Y., Lai, Z., Church, S., Jackson, L., and Bradford, K.
 lettuce and sunflower ESTs from the Compositae Genome Project
 Unpublished (2002)
<http://compenomics.ucdavis.edu/>

TITLE

JOURNAL

COMMENT

Contact: Alexander Kozik [R.W.Micheltore]
 Department of Vegetable Crops, R.W.Micheltore Lab
 University of California at Davis (UCD)
 Asmudson Hall, UCD, Davis, CA 95616, USA
 Tel: 1-(530)-742-1742
 Fax: 1-(530)-752-9659
 Email: akozike@ucdavis.edu [micheltore@vegmail.ucdavis.edu]
 belongs to contig OG_CA_Contig1411, see <http://cgdb.ucdavis.edu/>
 for details.
 Plate: QGB28 row: L column: 15.

FEATURES

source

Location/Qualifiers
 1. .755
 /organism="Lactuca sativa"
 /cultivar="Salinas"
 /db_xref="taxon:4236"
 /clone="QGB28L15"
 /clone_lib="QG-ABCDI lettuce salinas"
 /lab_host="E.coli"

/note="Vector: pBRCDNASfiAB; The library was constructed from 10 different sources of RNA from a single genotype. Separate cDNAs were generated using primers that incorporated unique 5' and 3' tags to distinguish each source of RNA. cDNAs were then pooled, size-fractionated, directionally cloned into a custom medium-copy vector and transformations made with four size classes to minimize size bias. Details of each source of RNA and library construction can be obtained at <http://cgdb.ucdavis.edu/>
 TAG_LIB-QG-ABCDI lettuce salinas
 TAG_TISSUE=flowers post-fertilized
 TAG_SEQ=TCGCATCGG"

BASE COUNT 223 a 149 c 159 g 223 t 1 others
 ORIGIN

Alignment Scores:

Pred. No.: 4,04e-124 Length: 755
 Score: 1048.50 Matches: 194
 Percent Similarity: 88.49% Conservative: 29
 Best Local Similarity: 76.98% Mismatches: 28
 Query Match: 38.22% Indels: 1
 DB: 14 Gaps: 1

US-09-701-395A-23 (1-529) x B0856148 (1-755)

QY 256 GlyGlyProArgValCysValGlnThrAlaTyrGlyValGluValGluAsnAsn 275
 Db 1 GGGGGTCCCGTGTGGTCCAAACAGCTTATGGTAGAGGTGAGGTGAAACAAAC 60
 QY 276 ProTyrAspProAsnLeuMetValPheMetAspTyrArgAspTyrMetGlnGlnLysLeu 295
 Db 61 CCTATGATCCAGATCAATTAAGGGTTCATGGATTATAGAGCTCTCAAAACATAAACCG 120
 QY GlnCysSerGluGluGluTyrProThrPheLeuTyrValMetProMetSerProThrArg 315
 Db 121 GAATCTTTAGAACCAAAATATCCGACTTCTCTATGTCATGGCCATGCTCCACACAAA 180
 QY 316 LeuPhePheGluGluThrCysLeuAlaSerLysAspAlaMetProPheAspLeuLys 335
 Db 181 ATATCTCTTTGAGGAACTTCTAGCTTCAAGGAGCCATGCTTCAATCTCTAAAG 240
 QY 336 ArgLysLeuMetSerArgLeuLysThrLeuGlyIleGlnValThrLysValTyrGluGlu 355
 Db 241 TCCAAACTCATGTCACGATTAAAGGCAATGGGTATCCGAATAACAGACGATAGGAAG 300
 QY 356 GluTrpSerTyrIleProValGlyGlySerLeuProAsnThrGluGlnLysAsnLeuAla 375
 Db 301 GAATGGTTCGTATATCCCGCTAGGTGGATCGTACCTAATACAGACAAAGATCTCGCA 360

QY 376 PheGlyAlaAlaAlaSerMetValHisProAlaThrGlyTyrSerValValArgSerLeu 395
 Db 361 TTTGGTCTGCAGCTAGTATGGTGCACCTGGCCACAGGTATTCAGTGTTCGATCTTTG 420
 QY 396 SerGluAlaProLysTyrAlaSerValIleAlaLysIleLeuLysGlnAspAsnSerAla 415
 Db 421 TCAGAAGCTCTAATATATGCAGCAGTCATTCGTAAGATTTTAAGACAGATCAATCTAAA 480
 QY 416 TyrValValSerGlyGlnSerSerAlaValAsnIleSerMetGlnAlaTtpSerSerLeu 435
 Db 481 GAAATGATTCTCT--CTTGGAAATACACATTAATTTCAAACAGCATGGGAACATTTG 537
 QY 436 TrpProLysGluArgLysArgGlnArgAlaPhePheLeuPheGlyLeuGluLeuVal 455
 Db 538 TGCCACTTCAAGGAAAGACAGCAGCCTCTCTTCTATTCGGACTATCACACATCGT 597
 QY 456 GlnLeuAspIleGluAlaThrArgThrPhePheArgThrPhePheArgLeuProThrTyr 475
 Db 598 CTAATGGATCTAGAGGGAACACGTCATTTTTCGGTACTTTCTTCGTTTCGCCANATGG 657
 QY 476 MetTrpTrpGlyPheLeuGlySerSerLeuSerSerPheAspLeuValLeuPheSerMet 495
 Db 658 ATGTGGTGGGATTTTGGGGTCTCTTCTTATCTCAACGGATTGTGATAATATTTCGGCTT 717
 QY 496 TyrMetPheValLeuAlaProAsnSerMetArgMet 507
 Db 718 TATATGTTTGTATAGCACCCTCACAGCTTGAGATG 753

RESULT 2

BQ623342

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BQ623342 664 bp mRNA linear EST 01-JUL-2002
 USDA-FP_00433 Ridge pineapple sweet orange entire seedling Citrus
 sinensis cDNA clone USDA-FP_00433 5', mRNA sequence.

USDA-FP_00433.1 GI:21650511

Citrus sinensis.

Citrus sinensis.

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eurosids II; Sapindales; Rutaceae; Citrus.

1 (bases 1 to 664)

Bausher, M., McKendree, W., Dang, P., Chaparro, J., Shatters, R., Hunter

, W. and Niedz, R.

Expressed sequence tags isolated from entire sweet orange (C.

sinensis L. Osbeck) seedling

Unpublished (2003)

Contact: Michael Bausher

US Horticultural Research

USDA - ARS

2001 South Rock Rd., Fort Pierce, FL 34945, USA

Tel: (772) 462-5918

Fax: (772) 462-5961

Email: mbausher@ushrl.ars.usda.gov

Seq primer: T3 Primer.

Location/Qualifiers

1. .664

/organism="Citrus sinensis"

/cultivar="Ridge Pineapple"

/db_xref="taxon:2711"

/clone="USDA-FP_00433"

/clone_lib="Ridge pineapple sweet orange entire seedling"

/tissue_type="entire seedling"

/dev_stage="50 days after germination"

/lab_host="XL1-Blue"

/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:

XhoI; A high quality EST with at least 200 contiguous

bases at Trace runner score of 20 or better"

BASE COUNT 180 a 109 c 177 g 198 t

ORIGIN

Alignment Scores:


```

Db 599 AGTATTCACTCAAGTGTGAACACCTTTGGCCCAAGAACGAAACAGACATAGATCG 658
:::||||| |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
RESULT 4
BQ994803 661 bp mRNA linear EST 21-AUG-2002
LOCUS QGF8D16.yg.ab1 QG_EFGHJ lettuce serriola Lactuca sativa cDNA clone
DEFINITION QGF8D16, mRNA sequence.
ACCESSION BQ994803
VERSION BQ994803.1 GI:22414338
KEYWORDS EST.
SOURCE Lactuca sativa.
ORGANISM Lactuca sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids II; Asterales; Asteraceae; Lactuceae;
Lactuca.
REFERENCE 1 (bases 1 to 661)
AUTHORS Kozik,A., Michelmore,R.W., Knapp,S., Matvienko,M., Rieseberg,L.,
Lin,H., van Damme,M., Lavelle,D., Chevalier,P., Ziegler,J., Ellison
P., Kolkman,J., Slabaugh,M.S., Livingston,K., Zhou,Y., Lai,Z.,
Church,S., Jackson,L. and Bradford,K.
TITLE Lettuce and Sunflower ESTs from the Compositae Genome Project
JOURNAL http://compgenomics.ucdavis.edu/
COMMENT Unpublished (2002)
Contact: Alexander Kozik [R.W.Michelmore]
Department of Vegetable Crops, R.W.Michelmore Lab
University of California at Davis (UCD)
Asmundson Hall, UCD, Davis, CA 95616, USA
Tel: 1-(530)-742-1742
Fax: 1-(530)-752-9659
Email: akozik@ucdavis.org [michelmore@vegmall.ucdavis.edu]
belongs to contig QG_CA_Contig1411, see http://cgpdb.ucdavis.edu/
for details.
Plate: QGF8 row: D column: 16.
FEATURES
source
Location/Qualifiers
1..861
/organism="Lactuca sativa"
/cultivar="L.serriola"
/db_xref="taxon:4236"
/clone="QGF8D16"
/clone_lib="QG_EFGHJ lettuce serriola"
/lab_host="E.coli"
/note="Vector: pBRCDNASFIAB; The library was constructed
from 10 different sources of RNA from a single genotype.
Separate cDNAs were generated using primers that
incorporated unique 5' and 3' tags to distinguish each
source of RNA. cDNAs were then pooled, size-fractionated,
directionally cloned into a custom medium-copy vector and
transformations made with four size classes to minimize
size bias. Details of each source of RNA and library
construction can be obtained at http://cgpdb.ucdavis.edu/
TAG_LIB=QG_EFGHJ lettuce serriola
TAG_TISSUE=flowers pre-fertilized
TAG_SEQ=CGTTGACGGG"
BASE COUNT 196 a 133 c 135 g 197 t
ORIGIN
Alignment Scores:
Pred. No.: 7 49e-100 Length: 661
Score: 860.00 Matches: 167
Percent Similarity: 85.71% Conservative: 25
Best Local Similarity: 74.55% Mismatches: 28
Query Match: 31.35% Indels: 5
DB: 14 Gaps: 2
US-09-701-395A-23 (1-529) x BQ994803 (1-661)
QY 294 LysLeuGlnCysSerGluGluThrProThrPheLeuTyrrValMetProMetSerPro 313
|||::: |||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 3 AAACCGGAATCTTTAGAACGAAATATCCGACTTCTCTATGTCATGCAATGCTCCA 62
QY 314 ThrArgLeuPheGluThrCysLeuAlaSerLysAspAlaMetPropheAspLeu 333

```

```

Db 63 ACAAATAATCTCTCGAGGAACCTTGTTAGCTTCAAGAGAACCCATGCTTCAATCTT 122
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 334 LeuLysArgLysLeuMetSerArgLeuLysThrLeuGlyIleGlnValThrLysValTy 353
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 123 CTAAAGTCAAACTCATGTCACGATTAAGGCAATGGGTATCCGATAACAAGAAC-- 178
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 354 GluGluLysTrpSerTyrrIleProValGlyGlySerLeuProAsnThrGluGlnLysAsn 373
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 179 -----GAATGGTCGTATATATCCCGTAGGTGGATCGTTACCTAATACAGAAAGAAT 232
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 374 LeuAlaPheGlyAlaAlaAlaSerMetValHisProAlaThrGlyTyrrSerValValArg 393
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 233 CTGGCAATTTGGTCTCGAGCTAGCATGGTCCACCCCTCCACAGGATTCAGTTGCCGA 292
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 394 SerLeuSerGluAlaProLysTyrrAlaSerValIleAlaLysIleLeuLysGlnAspAsn 413
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 293 TCTTTGTCAGAACTCTTAATATGACAGCATCTATGCTAAGATTTTAACACAGATCAA 352
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 414 SerAlaTyrrValValSerGlyGlnSerSerAlaValAsnIleSerMetGlnAlaTrpSer 433
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 353 TCTAAAGAGATGATTTCT---CTTGAAATAACATTAACATTTCAAACAAGCATGGAA 409
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 434 SerLeuTrpProLysGluArgLysArgGlnArgAlaPhePheLeuPheGlyLeuLysLeu 453
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 410 ACATTGGCCACTTGAAGGAAAGACACGAGCCCTCTTCTGTTCCGACTATCACAC 469
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 454 IleValGlnLeuAspIleGluAlaThrArgThrPhePheArgThrPhePheArgLeuPro 473
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 470 ATCGTCTAATGATCTAGAGGGAACACGATACATTTTCCGCTACTTCTTCTGTTGCC 529
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 474 ThrTrpMetTrpTrpGlyPheLeuGlySerSerLeuSerSerPheAspLeuValLeuPhe 493
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 530 AAATGATGATGGTGGGATTTTGGGCTCTCTTTATCTTCAACGGGATTTGATAATATT 589
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 494 SerMetTyrrMetPheValLeuAlaProAsnSerMetArgMetSerLeuValArgHisLeu 513
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 590 GCGCTTATATGTTGTGTAGTACACCTCACAGCTTGAGATGAAGTGGGTAGACATCTA 649
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 514 LeuSerAspPro 517
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 650 CTTTCTGATCCG 661
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
RESULT 5
AV931085 719 bp mRNA linear EST 18-JAN-2002
LOCUS AV931085 K. Sato unpublished cDNA library, cv. Haruna Nijo second
DEFINITION leaf stage seedling leaves Hordeum vulgare subsp. vulgare cDNA
clone basd18h18 3', mRNA sequence.
ACCESSION AV931085
VERSION AV931085.1 GI:18226882
KEYWORDS EST.
SOURCE Hordeum vulgare subsp. vulgare.
ORGANISM Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
; Triticeae; Hordeum.
REFERENCE 1 (bases 1 to 719)
AUTHORS Sato,K., Saisho,D. and Takeda,K.
TITLE Barley EST sequencing project in NIG and Okayama Univ
JOURNAL Unpublished (2002)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
1..719
/organism="Hordeum vulgare subsp. vulgare"
/cultivar="Haruna Nijo"
/db_xref="taxon:112509"
FEATURES
source

```

/clone="basd1818"
 /clone_lib="K. Sato unpublished cDNA library, cv. Haruna
 Nijo second leaf stage seedling leaves"
 /tissue_type="seedling leaves"
 /dev_stage="second leaf stage"
 BASE COUNT 207 a 158 c 170 g 182 t 2 others
 ORIGIN

Alignment Scores:
 Pred. No.: 1.18e-97 Length: 719
 Score: 843.50 Matches: 164
 Percent Similarity: 84.9% Conservative: 28
 Best Local Similarity: 72.5% Mismatches: 33
 Query Match: 30.7% Indels: 2
 DB: 10 Gaps: 1

US-09-701-395A-23 (1-529) x AV931085 (1-719)

QY 303 ProThrPheLeuTyrValMetProMetSerProThrArgLeuPheGluThrCys 322
 |||||
 Db 713 CCAACATTTCTTATGTCATGCCATGTCATCCACACGAGTTTCTTTGAGGAACATGC 654
 QY 323 LeuAlaSerLysAspAlaMetProPheAspLeuLeuLysArgLysLeuMetSerArgLeu 342
 |||||
 Db 653 TTAGCTTCAAGAGTCAATGCGCTTGTATCTCTTAAGAGAGGTGTGATGCTCGGTTG 594
 QY 343 LysThrLeuGlyIleGlnValThrLysValTyrGluGluThrPyrTrpSerValProVal 362
 |||||
 Db 593 GATGCGATGGGATCGTATCTTAAAGTATACGAGGAGGTGTGTTATATCTCTGTT 534
 QY 363 GlyGlySerLeuProAsnThrGluGlnLysAsnLeuAlaPheGlyAlaAlaSerMet 382
 |||||
 Db 533 GGAGGATCCTTACCTAACACACACAGCAAAATCTTGCATTTGGTGTGCGAGGATG 474
 QY 383 ValHisProAlaThrGlyTyrSerValValArgSerLeuSerGluAlaProLysTyrAla 402
 |||||
 Db 473 GTCCACCTCGCACTGGATGTCGTTGTCAGATCTTGTCTGAAGCTCAAGATATGCT 414
 QY 403 SerValIleAlaLysIleLeu---LysGlnAspAsnSerAlaTyrValValSerGlyGln 421
 |||||
 Db 413 TCTGTGATATCTGATATCTTAGAAATCGTCTATCTTGGCAATATTTGCTTGAAGT 354
 QY 422 SerSerAlaValAsnIleSerMetGlnAlaTrpSerSerLeuTrpProLysGluArgLys 441
 |||||
 Db 353 TCTGAATGTCAGTCCATCAATGCTGATGGGAAACACTATGGCTCAAGAACGGAAA 294
 QY 442 ArgGlnArgAlaPhePheLeuPheGlyLeuGluLeuLeuValGlnLeuAspIleGluAla 461
 |||||
 Db 293 CGTCAACGCTATCTTCTCTTGGATTGGCTTGATTAATCAACTGGATACGAGGC 234
 QY 462 ThrArgThrPheArgThrPheArgLeuProThrTrpMetTrpTrpGlyPheLeu 481
 |||||
 Db 233 ATCCAAACGTTCTCGAAACCTTTTCCGGTTACCAATGGATGGGCGAGGATTCCTT 174
 QY 482 GlySerSerLeuSerSerPheAspLeuValLeuPheSerMetTyrMetPheValLeuAla 501
 |||||
 Db 173 GGTTCGACCTGCTGTCAGTGTGATCTGCTGCTTGTGCACTTACATGTTTGAATGG 114
 QY 502 ProAsnSerMetArgMetSerLeuValArgHisLeuLeuSerAspProSerGlyAlaVal 521
 |||||
 Db 113 CCAACAAATTTGCGAATGAATCGTGCAGACACCTCTCTCGGACCCAC -GGTTCAGCA 55
 QY 522 MetValArgAlaTyrLeu 527
 |||||
 Db 54 ATGATCAGGACCTACCTG 37

RESULT 6
 AV926200
 LOCUS
 DEFINITION AV926200 K. Sato unpublished cDNA library, cv. Haruna Nijo second
 leaf stage seedling leaves Hordeum vulgare subsp. vulgare cDNA
 clone basd1818 5', mRNA sequence.
 AV926200
 AV926200

VERSION
 KEYWORDS
 SOURCE
 ORGANISM

AV926200.1 GI:18221997

EST.
 Hordeum vulgare subsp. vulgare.
 Hordeum vulgare subsp. vulgare
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
 ; Triticeae; Hordeum.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

1 (bases 1 to 665)
 Sato, K., Saisho, D. and Takeda, K.
 Barley EST sequencing project in NIG and Okayama Univ
 Unpublished (2002)
 Contact: Tadasi Shin-i
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshin@genes.nig.ac.jp.

Location/Qualifiers
 1. 665
 /organism="Hordeum vulgare subsp. vulgare"
 /cultivar="Haruna Nijo"
 /db_xref="taxon:112509"
 /clone="basd1818"
 /clone_lib="K. Sato unpublished cDNA library, cv. Haruna
 Nijo second leaf stage seedling leaves"
 /tissue_type="seedling leaves"
 /dev_stage="second leaf stage"
 BASE COUNT 170 a 145 c 151 g 197 t 2 others
 ORIGIN

FEATURES
 source

1. 665
 /organism="Hordeum vulgare subsp. vulgare"
 /cultivar="Haruna Nijo"
 /db_xref="taxon:112509"
 /clone="basd1818"
 /clone_lib="K. Sato unpublished cDNA library, cv. Haruna
 Nijo second leaf stage seedling leaves"
 /tissue_type="seedling leaves"
 /dev_stage="second leaf stage"
 BASE COUNT 170 a 145 c 151 g 197 t 2 others
 ORIGIN

Alignment Scores:
 Pred. No.: 2.99e-91 Length: 665
 Score: 793.50 Matches: 153
 Percent Similarity: 81.36% Conservative: 26
 Best Local Similarity: 69.5% Mismatches: 40
 Query Match: 28.93% Indels: 1
 DB: 10 Gaps: 1

US-09-701-395A-23 (1-529) x AV926200 (1-665)

QY 288 ArgAspTyrMetGlnLysLeuGlnCysSerGluGluGluTyrProThrPheLeuTyr 307
 |||||
 Db 4 AGATTGTTTCAAGNAGNAGTTCTCACACCTTGAGGAGCCCAATCCACATTTCTTAT 63
 QY 308 ValMetProMetSerProThrArgLeuPheGluGluThrCysLeuAlaSerLysAsp 327
 |||||
 Db 64 GCATGGCCTGTCATCCACGAGTTTCTTTGAGGAACATGCTTAGCTTCAAAAGAT 123
 QY 328 AlaMetProPheAspLeuLeuLysArgLysLeuMetSerArgLeuLysThrLeuGlyIle 347
 |||||
 Db 124 GCAATGCCCTTTGATCTCTTAAGAGAGGTGATGTCCTCGTTGGATGGGAGTC 183
 QY 348 GlnValThrLysValTyrGluGluGluTrpSerTyrIleProValGlySerLeuPro 367
 |||||
 Db 184 CGTATCTAAAGATATACGAGGAGGTGCTTATATTTCTCTTGGAGGATCCTTACCT 243
 QY 368 AsnThrGluGlnLysAsnLeuAlaPheGlyAlaAlaSerMetValHisProAlaThr 387
 |||||
 Db 244 ACACAGACAGCAAAATCTTGCATTTGGTGTGCGAGGATGTTCCACCTGCAACT 303
 QY 388 GlyTyrSerValValArgSerLeuSerGluAlaProLysTyrAlaSerValIleAlaLys 407
 |||||
 Db 304 GGATACCTCGTGTGATCTTGTCTGAAGCTTCAAGATATGCTTCTGTGATATCTGAT 363
 QY 408 IleLeu---LysGlnAspAsnSerAlaTyrValValSerGlyGlnSerSerAlaValAsn 426
 |||||
 Db 364 ATCTTAGGAAATCGTCTCTATTCTGGCAATATTTGCTTGAAGTCTTGAATGTCAGT 423
 QY 427 IleSerMetGlnAlaTrpSerLeuTrpProLysGluArgLysArgGlnArgAlaPhe 446
 |||||
 Db 424 CCATCAATGCTTGCATGGGAAACACTATGGCTTCAAGACGGAACGCTCAACGCTCATTC 483

```

QY 447 PheLeuPheGlyLeuGluLeuValGlnLeuAspIleGluAlaThrArgThrPhePhe 466
Db 484 TTCCTCTTTGGATGGCTTGATAATTAACCTGGAACGACGATCCAAACGTTCTTC 543
QY 467 ArgThrPhePheArgLeuProThrTyrMetTyrPheGlyPheLeuGlySerSerLeuSer 486
Db 544 GAACCTTTTCCGGTACCAATGATGATGGGAGAGATTCCTTGGTTCACCGCTGTCG 603
QY 487 SerPheAspLeuValLeuPheSerMetTyrMetPheValLeuAlaProAsnSerMetArg 506
Db 604 TCAGTGGATCTCATGCTGTGTTCACCTCTACATGTTTGCATTAATGCGCCCAACAAATTGCGA 663
RESULT 7
BOL16144 610 bp mRNA linear EST 22-JUL-2002
LOCUS EST601720 mixed potato tissues Solanum tuberosum cDNA clone STMDG03
5' end, mRNA sequence.
ACCESSION BOL16144
VERSION BOL16144.2 GI:21917328
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
1 (bases 1 to 610)
Buell,C.R., Hart,A., Baker,B., Tanksley,S., Fry,W., Smart,C.,
Restrepo,S., Griffiths,H., van der Hoeven,R., Tsai,J. and
Karanycheva,S.A.
Generation of a set of potato cDNA clones for microarray analyses
Unpublished (2002)
On Apr 17, 2002 this sequence version replaced gi:20168106.
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato@igr.org
This clone is available through the Research Genetics, contact the
Research Genetics for further information 1-800-711-6195 or
cdna@resgen.com
Seq primer: T3.
FEATURES
source Location/Qualifiers
1. .610
/organism="Solanum tuberosum"
/cultivar="Kennebec or Binjete"
/db_xref="taxon:4113"
/clone="STMDG03"
/clone_lib="mixed potato tissues"
/tissue_type="mixed tissues"
/lab_host="SOLR"
/notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: Combination of untreated and Phytophthora
infestans-treated libraries of stolons, leaves, leaflets,
axillary buds of stem explants, petioles, germinating eyes
tubers, or roots."
BASE COUNT 187 a 128 c 127 g 168 t
ORIGIN
Alignment Scores:
Pred. No.: 6,51e-83 Length: 610
Score: 728.50 Matches: 151
Percent Similarity: 74.89% Conservative: 25
Best Local Similarity: 64.26% Mismatches: 26
Query Match: 26.56% Indels: 35
DB: 14 Gaps: 2
US-09-701-395a-23 (1-529) x BOL16144 (1-610)
QY 206 ValSerTyrLeuAspSerLysValGluArgIleThrGluAlaGlyAspGlyHisSerLeu 225
Db 2 GTTCTGTATCTAACTCAAGAGTGGATAGTATTGTGAGGCACAAATGCCACATCTT 61
QY 226 ValValCysGluAsnGluIlePheIleProCysArgLeuAlaThrValAlaSerGlyAla 245

```

```

Db 62 GT----- 63
QY 246 AlaSerGlyLysLeuLeuGluTyrGluValGlyProArgValCysValGlnThrAla 265
Db 64 -----AGAGTTTCTGTTCAACACGCT 84
QY 266 TyrGlyValGluValGluValGluAsnAsnProTyrAspProAsnLeuMetValPheMet 285
Db 85 TATGGAGTGAAGTTGAGTTGATACCAATCATATGACCGACGCTGATGGTTTTCATG 144
QY 286 AspTyrArgAspTyrMetGlnGlnLysLeuGlnCysSerGluGluGluTyrProThrPhe 305
Db 145 GATTATAGAGACTATGTTCAGACGACGCTCAATCTTTAGAAAGTAATAATCCAACTTT 204
QY 306 LeuTyrValMetProMetSerProThrArgLeuPheGluGluThrCysLeuAlaSer 325
Db 205 CTCATGCCATGCCCATGTCTCCAAACAGAGTCTTTTCGAGGAAACTTTGTTGGCTTCA 264
QY 326 LysAspAlaMetProPheAspLeuLeuLysArgLysLeuMetSerArgLeuLysThrLeu 345
Db 265 AAAGATGCAATGCCATTCGATCTGTAAAGAAAAAACTGATGTTCAGATTGAACACCTC 324
QY 346 GlyIleGlnValThrLysValTyrGluGluTyrSerTyrIleProValGlyGlySer 365
Db 325 GGTGTAGAATTAAGAAATTTACGAGGAGGAATGGTCTTACATACACAGTTGGAGGATCT 384
QY 366 LeuProAsnThrGluGlnLysAsnLeuAlaPheGlyAlaAlaAlaSerMetValHisPro 385
Db 385 TTGCCAAATACAGAACAAAAACATTCGATTTGGTCTGCTGCTACATGTTGTTTCATCCA 444
QY 386 AlaThrGlyTyrSerValValArgSerLeuSerGluAlaProLysTyrAlaSerValIle 405
Db 445 GCCACAGGTTATTTCAGTCTGCATCAGTCTCTGCAAGCTCCAAATCGCGCTCGTCT 504
QY 406 AlaLysLeuLysGlnAspAsnSerAlaTyrValValSerGlyGlnSerSerAlaVal 425
Db 505 GGCATATATTACGACAAATATATAGCAAGACATGCTTACT---AGTTCAAGTACCCTG 561
QY 426 AsnIleSerMetGlnAlaTrpSerSerLeuTrpProLysGluArg 440
Db 562 AGTATTTCAACTCAAGCTTGGAC-ACCTCTTGGCCACAGAAGCA 605
RESULT 8
LOCUS BG241188 541 bp mRNA linear EST 15-FEB-2001
DEFINITION OVI_39_C08_b1_A002 Ovary 1 (OVI), Sorghum bicolor cDNA, mRNA
sequence.
ACCESSION BG241188
VERSION BG241188.1 GI:12776261
KEYWORDS EST.
SOURCE sorghum.
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 541)
Cordonnier-Pratt,M.-M., Gingle,A., Marsala,C., Sudman,M. and Pratt
,L.H.
An EST database from Sorghum: ovaries of varying immature stages
Unpublished (2000)
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions
below Phred quality 16. The threshold for highest quality sequence
is 20.
Seq primer: JEN REV
High quality sequence stop: 526
POLYA-No.

```

FEATURES source Location/Qualifiers
 1..541
 /organism="Sorghum bicolor"
 /db_xref="taxon:4558"
 /clone_lib="Ovary 1 (OVI)"
 /note="Organ: Mix of ovaries of varying immature stages from 8-week-old plants; Vector: pBluescript II from Lambda Zap II; Site_1: XhoI; Site_2: EcoRI; The library was made from poly-A RNA in the cloning vector lambda ZAP II. Clones to be sequenced were prepared by mass excision."
 BASE COUNT 146 a 107 c 140 g 147 t 1 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.9e-79 Length: 541
 Score: 701.00 Matches: 131
 Percent Similarity: 85.47% Conservative: 22
 Best Local Similarity: 73.18% Mismatches: 26
 Query Match: 25.56% Indels: 0
 DB: 12 Gaps: 0
 US-09-701-395A-23 (1-529) x BG241188 (1-541)
 Qy 159 GluArgCysIleGluHisAlaTrpLysAspThrIleValTyrLeuAspAsnAspAlaPro 178
 Db 5 GAGGTTGTATCGACACGCTCTGGAAGGATACATTGTCTACCTAGACAAATAACAGGCC 64
 Qy 179 ValLeuIleGlyArgAlaTyrGlyArgValSerArgHisLeuLeuHisGluGluLeuLeu 198
 Db 65 ATACTGATGGCGTCTTATGGCAGGTGCACCGTGACTGCTCCATGAGGAGTACTG 124
 Qy 199 LysArgCysValGluSerGlyValSerTyrLeuAspSerLysValGluArgIleThrGlu 218
 Db 125 AGAAGATGCTATGAAGCTGGGTGCACATCTGAACTCCAAAGTGCACAAATCATAGAA 184
 Qy 219 AlaGlyAspGlyHisSerLeuValValCysGluAsnGluIlePheIleProCysArgLeu 238
 Db 185 TCTCAGATGGACACAGATAGTCTGTGTATAGAGGTGCTGAGATATATTGACAGGCTT 244
 Qy 239 AlaThrValAlaSerGlyAlaAlaSerGlyLysLeuLeuGluTyrGluValGlyGlyPro 258
 Db 245 GCCATTGTGCTCGGGGGCAGCATCTGGTAGGCTCTAGAGTATGAGTGGGGGTCT 304
 Qy 259 ArgValCysValGlnThrAlaTyrGlyValGluValGluValGluAsnProTyrAsp 278
 Db 305 CGTGTTCGTCGACACTGCATCGGAGTAGAAGTTGAGGTGGAAATAATCCATATGAT 364
 Qy 279 ProAsnLeuMetValPheMetAspTyrArgAspTyrMetGlnGlnLysLeuGlnCysSer 298
 Db 365 CCCAGCTTAATGGTTTTCATGGACTACAGAGATTGTTTCAAGAGGAATCTCACACACT 424
 Qy 299 GluGluGluTyrProThrPheLeuTyrValMetProMetSerProThrArgLeuPhePhe 318
 Db 425 GAACAAGAAATCCACATCTCTGTATGCTATGCCATGTCTATCCACAGAGTTNTCTT 484
 Qy 319 GluGluThrCysLeuAlaSerLysAspAlaMetPropheAspLeuLysArgLys 337
 Db 485 GAGGAACATGCTAGCTTCTAAAGATGCTATGCTCTTGTATCTTACTTAAAGAAGAGG 541
 RESULT 9
 AW906884
 LOCUS EST343007 potato stolon, Cornell University Solanum tuberosum cDNA
 DEFINITION
 clone cSTA4H14, mRNA sequence.
 ACCESSION AW906884
 VERSION AW906884.1 GI:8071094
 KEYWORDS EST.
 SOURCE potato.
 ORGANISM Solanum tuberosum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 REFERENCE 1 (bases 1 to 548)

AUTHORS van der Hoeven, R., Bezzerides, J., Bachem, C., Horvath, B., Visser, R., Holt, I. E., Liang, F., Hansen, T. S., Utterback, T., Bowman, C. L., Doan, B., Bougri, O., Buell, C. R., Ronning, C. M., Tanksley, S. D. and Baker, B.
 TITLE Generation of ESTs from potato swelling stolons
 JOURNAL Unpublished (1999)
 COMMENT Contact: Research Genetics, Libraries Division
 Tel: 1-800-711-6195
 Email: cdna@resgen.com
 5 prime sequence.
 FEATURES Location/Qualifiers
 source 1..548
 /organism="Solanum tuberosum"
 /cultivar="Bintje"
 /db_xref="taxon:4113"
 /clone="cSTA4H14"
 /clone_lib="potato stolon, Cornell University"
 /tissue_type="axillary buds of stem explants, swelling stolons"
 /dev_stage="1 to 3 days"
 /lab_host="SOLR"
 /note="Vector: pBlueScript SK(-); Site_1: EcoRI; Site_2: XhoI; RNA was supplied by Christian Bachem & Beatrix Horvath (Laboratory of Plant Breeding, Dept. of Plant Sciences, Wageningen University, The Netherlands). Total RNA was isolated from developing axillary buds of potato nodal stem cuttings cultured on medium for the introduction of tuber formation as described in Bachem et al. (Plant Journal 1996). Tissue samples were taken of stages corresponding to growing stolons and the early stages of tuber formation."
 BASE COUNT 169 a 111 c 114 g 154 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.16e-78 Length: 548
 Score: 695.00 Matches: 139
 Percent Similarity: 76.08% Conservative: 20
 Best Local Similarity: 66.51% Mismatches: 18
 Query Match: 25.34% Indels: 33
 DB: 10 Gaps: 1
 US-09-701-395A-23 (1-529) x AW906884 (1-548)
 Qy 206 ValSerTyrLeuAspSerLysValGluArgIleThrGluAlaGlyAspGlyHisSerLeu 225
 Db 2 GTTCTGTATCTAAACTCGAAGTGGATGAGATTGTGAGCCACAAATGCCACAGCTT 61
 Qy 226 ValValCysGluAsnGluIlePheIleProCysArgLeuAlaThrValAlaSerGlyAla 245
 Db 62 GT----- 63
 Qy 246 AlaSerGlyLysLeuLeuGluTyrGluValGlyGlyProArgValCysValGlnThrAla 265
 Db 64 -----AGAGTTCTCTTCAACACAGCT 84
 Qy 266 TyrGlyValGluValGluValGluAsnProTyrAspProAsnLeuMetValPheMet 285
 Db 85 TATGAGTGGAGTTGAGTTGATACCAATCCATATGACCCGAGCTGATGTTTCATG 144
 Qy 286 AspTyrArgAspTyrMetGlnGlnLysLeuGlnCysSerGluGluGluTyrProThrPhe 305
 Db 145 GATTATAGAGATATGTCACACACGCTCAATCTTTAGAACGCTAAATATCCACATTT 204
 Qy 306 LeuTyrValMetProMetSerProThrArgLeuPheGluGluThrCysLeuAlaSer 325
 Db 205 CTCATGCCATGCCATGCTCCACACAGGCTTTTCGAGGAACACTTGTTCCTTCA 264
 Qy 326 LysAspAlaMetPropheAspLeuLeuLysArgLysLeuMetSerArgLeuLysThrLeu 345
 Db 265 AAAGATGCAATGCCATTCGATCTGTTAAAGAAAAAACTGATGTTAGGATTGAACACCTC 324
 Qy 346 GlyIleGlnValThrLysValTyrGluGluGluTyrSerTyrIleProValGlyGlySer 365

van der Hoeven, R., Bezzerides, J., Bachem, C., Horvath, B., Visser, R., Holt, I. E., Liang, F., Hansen, T. S., Utterback, T., Bowman, C. L., Doan, B., Bougri, O., Buell, C. R., Ronning, C. M., Tanksley, S. D. and Baker, B.
 Generation of ESTs from potato swelling stolons
 Unpublished (1999)
 Contact: Research Genetics, Libraries Division
 Tel: 1-800-711-6195
 Email: cdna@resgen.com
 5 prime sequence.
 Location/Qualifiers
 source 1..548
 /organism="Solanum tuberosum"
 /cultivar="Bintje"
 /db_xref="taxon:4113"
 /clone="cSTA4H14"
 /clone_lib="potato stolon, Cornell University"
 /tissue_type="axillary buds of stem explants, swelling stolons"
 /dev_stage="1 to 3 days"
 /lab_host="SOLR"
 /note="Vector: pBlueScript SK(-); Site_1: EcoRI; Site_2: XhoI; RNA was supplied by Christian Bachem & Beatrix Horvath (Laboratory of Plant Breeding, Dept. of Plant Sciences, Wageningen University, The Netherlands). Total RNA was isolated from developing axillary buds of potato nodal stem cuttings cultured on medium for the introduction of tuber formation as described in Bachem et al. (Plant Journal 1996). Tissue samples were taken of stages corresponding to growing stolons and the early stages of tuber formation."


```

US-09-701-395A-23 (1-529) x AV521185 (1-614)
QY 364 GlySerLeuProAsnThrGluGlnLysAsnLeuAlaPheGlyValAlaAlaSerMetVal 383
Db 613 GGATCCTTACCTAACACAGACAGAAATCTGCATTTGGTGGCTGCAGGAGATGGTC 554
QY 384 HisProAlaThrGlyTyrSerValValArgSerLeuSerGluAlaProLysTyrAlaSer 403
Db 553 CACCCTGCAACCTGGATCTCGGTGGTCAGATCTTTGCTCTGAAGCTCCAAGATATGCTTCT 494
QY 404 ValIleAlaLysIleLeu---LysGlnAspAsnSerAlaTyrValValSerGlyGlnSer 422
Db 493 GTGATATCTGATATCTTAGAAATCGTCTATTCTGGACAATATTTGCTCGGAAGTTCT 434
QY 423 SerAlaValAsnIleSerMetGlnAlaTrpSerSerLeuTrpProLysGluArgLysArg 442
Db 433 GAATGTCAGTCCATCATGCTTGCATGGGGAACACTATGGCCCAAGAAACGCT 374
QY 443 GlnArgAlaPhePheLeuPheGlyLeuGluLeuIleValGlnLeuAspIleGluAlaThr 462
Db 373 CAACGCTCATCTCTCTTTGGATGGCTTGATAATCAACTGATGAACGAAGGCATC 314
QY 463 ArgThrPhePheArgThrPhePheArgLeuProThrTrpMetTrpTrpGlyPheLeuGly 482
Db 313 CAACGCTCTTCCGAACCTTTTCCGGTTACCCAAATGGATGTGGCGAGGATTCCTTGT 254
QY 483 SerSerLeuSerSerPheAspLeuValLeuPheSerMetTyrMetPheValLeuAlaPro 502
Db 253 TCGAGTGTCTCGTCAGTGTATCTCATGCTGTGGTGGCTCTACATGTTTCATTTGCCCA 194
QY 503 AsnSerMetArgMetSerLeuValArgHisLeuLeuSerAspProSerGlyAlaValMet 522
Db 193 AACAAATTCGAATGAACCTCGTCAGACACCTCTCTCGGACCCCACTGGTTCCAGCAATG 134
QY 523 ValArgAlaTyrLeu 527
Db 133 ATCAGACCTACCTG 119

RESULT 12
LOCUS AV5211661/c
DEFINITION AV5211661 Arabidopsis thaliana aboveground organs two to six-week old Arabidopsis thaliana cDNA clone AP264b02F 3', mRNA sequence.
ACCESSION AV5211661
VERSION AV5211661.1 GI:8681188
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 557)
AUTHORS Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
TITLE A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries
JOURNAL DNA Res. 7, 175-180 (2000)
MEDLINE 20363093
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
Location/Qualifiers
1, 557
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="AP264b02F"
/clone_lib="Arabidopsis thaliana aboveground organs two to six-week old"
/tissue_type="aboveground organs"

FEATURES
source
1, 557
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="AP264b02F"
/clone_lib="Arabidopsis thaliana aboveground organs two to six-week old"
/tissue_type="aboveground organs"

US-09-701-395A-23 (1-529) x AV5211661 (1-557)
QY 88 MetGluLysGlnAlaLysLeuAlaAspLysLeuProPheLeuProPheGlyGluSerVal 107
Db 552 ATGGATGAACAGCTCAAGCTTGTGATAAGTTGCCCTCATATATCAATTTGGTATGGTGCT 493
QY 108 MetAspLeuValIleLeuGlyCysGlyProAlaGlyLeuSerLeuAlaGluAlaAla 127
Db 492 TTGGATCTAGTGGTATTGGTGTCTGCTGCTGCTTGGCTTGGCTGAGATCAGCT 433
QY 128 LysLeuGlyLeuLysValGlyLeuIleGlyProAspLeuProPheThrAsnAsnTyrGly 147
Db 432 AAGCTTTGGATTAAAGTTGACATTCATTTGGTCCAGATCTTCTTTTACTAACAAATTACGGT 373
QY 148 ValTrpGluAspGluPheLysAspLeuGlyLeuGluArgCysIleGluHisAlaTrpLys 167
Db 372 GTTTGGGAAGATGAATTCATGATCTTGGGCTGCAAAATGATATTGAGCATGTTGGAGA 313
QY 168 AspThrIleValTyrLeuAspAsnAspAlaProValLeuIleGlyArgAlaTyrGlyArg 187
Db 312 GAGACTATTGTGTATCTGGATGATGACAAAGCTTATACCATTTGGCTGCTTATGAAGA 253
QY 188 ValSerArgHisLeuLeuHisGluGluLeuLeuLysArgCysValGluSerGlyValSer 207
Db 252 GTTAGTCGACGTTTGTCTCCATGAGGAGCTTTTGAGAGGCTGTCGAGTCAGGTCTCTCG 193
QY 208 TyrLeuAspSerLysValGluArgIleThrGluAlaGlyAspGlyHisSerLeuValVal 227
Db 192 TACCTTAGCTCGAAGATTGACACATACAGAAAGCTTCTGATGCCCTTAGACTTGTTTCA 133
QY 228 Cys 228
Db 132 TGT 130

RESULT 13
LOCUS AW982124
DEFINITION AW982124 Suaeda salsa ZAP cDNA library Suaeda maritima subsp. salsa cDNA similar to lycopene epsilon-cyclase, mRNA sequence.
ACCESSION AW982124
VERSION AW982124.1 GI:8173697
KEYWORDS EST.
SOURCE Suaeda maritima subsp. salsa.
ORGANISM Suaeda maritima subsp. salsa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Caryophyllales; Caryophyllales; Chenopodiaceae; Suaeda.
REFERENCE 1 (bases 1 to 465)
AUTHORS Zhang, L., Ma, C.L., Wang, P.P., Sun, Y.F., Zhao, Y.X. and Zhang, H.
TITLE Expressed sequence tags from a halophyte Suaeda salsa cDNA library
JOURNAL Unpublished (2000)
COMMENT Contact: Hui Zhang
Key Laboratory of Plant Stress Research
The Biology Department of Shandong Normal University
No.88, Wenhua East Road, Jinan, Shandong Province, 250014, PRC
Tel: (86)531-2960864
Fax: (86)531-2966954
Email: zhanghesdnu.edu.cn.
Location/Qualifiers

```

FEATURES

```

Source      1. .465
/organism="Suaeda maritima subsp. salsa"
/db_xref="taxon:126914"
/clone_lib="Suaeda salsa ZAP cDNA library"
/dev_stage="seedling"
/site="organ: aerial part tissue; Vector: lambda zap;
Site_1: EcORI; Site_2: XhoI; total RNA extraction from
NaCl(400mM) treated Suaeda salsa by RNAGENT kit(Promega);
mRNA isolation by MESSAGEMAKER kit(GIBCO BRL);
directional cDNA synthesis(EcORI XhoI) by cDNA synthesis
kit(STRATAGEN); the ZAP express library by GigapackIII
Gold Cloning kit(STRATAGENE)"
BASE COUNT 121 a 100 c 121 g 122 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 2.24e-66 Length: 465
Score: 599.00 Matches: 118
Percent Similarity: 88.51% Conservative: 13
Best Local Similarity: 79.73% Mismatches: 17
Query Match: 21.84% Indels: 1
DB: 1.0 Gaps: 0

US-09-701-395A-23 (1-529) x AW982124 (1-465)
QY 187 ArgValSerArgHisLeuLeuHisGluGluLeuLeuLysArgCysValGluSerGlyVal 206
DB 22 AGGTCACGCCGCAATTTACTTCAACAAGAGCTCGTCAAAAGGTGTTTCGAGTCAGGGGTC 81
QY 207 SerTyrLeuAspSerLysValGluArgIleThrGluAlaGlyAspGlyHisSerLeuVal 226
DB 82 TCTTATCTGAATGCAAAAGTCGATAGGATCATGGAAGAGGTGATGATCTAGTCTGTG 141
QY 227 ValCysGluAsnGluIlePheIleProCysArgLeuAlaThrValAlaSerGlyValAla 246
DB 142 GAAGA-GAAGATGATGATCACTATTTCCTGCAGGCTTCAACTGTTCATCTGGAGCAGCT 200
QY 247 SerGlyLysLeuLeuGluTyrGluValGlyGlyProArgValCysValGlnThrAlaTyr 266
DB 201 TCAGGGAAGCTCTGGAGTATGAAGTCCGGGGTCCAGGGTGTGTTCAACAGCCTAT 260
QY 267 GlyValGluValGluValGluAsnAspProTyrAspProAsnLeuMetValPheMetAsp 286
DB 261 GGTGTGAAGTCGAGGTGGAAACAGCTCTTATGATCCCAACGTGATGGTCTCATGGAC 320
QY 287 TyrArgAspTyrMetGlnGlnLysLeuGlnCysSerGluGluGluTyrProThrPheLeu 306
DB 321 TACAGATTACACCAACAGAGGTTTCAGTCTCCGGAGGCAAGTATCCAACTTCTTG 380
QY 307 TyrValMetProMetSerProThrArgLeuPhePheGluGluThrCysLeuAlaSerLys 326
DB 381 TACGCAATGCCACTGTCACTACTAGAACTCTTTCGAGGAGACTTGTCTGGCTTCAATA 440
QY 327 AspAlaMetProPheAspLeuLeu 334
DB 441 GATGCAATGCCATTTGATCTACTC 464

RESULT 14
AV822597
LOCUS AV822597 668 bp mRNA linear EST 01-APR-2002
DEFINITION AV822597 RAFL5 Arabidopsis thaliana cDNA clone RAFL05-09-E07 5',
mRNA sequence.
ACCESSION AV822597
VERSION AV822597.1 GI:19864646
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 668)
REFERENCE Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M.,
Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y.,

```

```

TITLE JOURNAL
COMMENT
Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y.
and Shinozaki, K.
Large scale analysis of Arabidopsis full-length cDNA (2002b)
Unpublished (2002)
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@tc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al., 1998). This clone is in a
modified pBluescript vector as a SstI/XhoI insert. Please visit our
web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for
further details.
FEATURES
Location/Qualifiers
source 1. .668
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/clone="RAFL05-09-E07"
/clone_lib="RAFL5"
/dev_stage="rosette plants"
/lab_host="SOLR"
/notes="Site_1: SstI; Site_2: XhoI; subjected to
dehydration-treated(1.2, 5, 10, 24 hr)"
BASE COUNT 175 a 95 c 183 g 214 t 1 others
ORIGIN

```

```

Alignment Scores:
Pred. No.: 4.88e-66 Length: 668
Score: 598.50 Matches: 120
Percent Similarity: 74.61% Conservative: 24
Best Local Similarity: 62.18% Mismatches: 37
Query Match: 21.82% Indels: 12
DB: 1.0 Gaps: 4

US-09-701-395A-23 (1-529) x AV822597 (1-668)
QY 1 MetGluLeuLeuGlyValArgAsnLeu-----IleSerSerCysProValTr 16
DB 106 ATGAGTGTGTGGGCTAGGAATTTCCGAGCAATGGCAGGTTTCAACATTTCCGTCATG 165
QY 16 pThrPheGlyThrArgAsnLeuSerSerLysLeuAlaTyrAsnIleHisArgTyrG1 36
DB 166 GAGTTCGGAAGAAATTTCCAGTGTGTTAAGAGATACACTATAGGAATATTCGTTTCGG 225
QY 36 ySerSerCysArgValAspPheGlnValArgAlaAspGlyGly-----SerGlySerAr 54
DB 226 T---TTGTGT-----AGTGTCAAGCTAGCGCGCGGGAAGTTCCGGTAGTGA 270
QY 54 gSerSerValAlaTyrLysGluGlyPheValAspGluGluAspPheIleLysAlaGlyG1 74
DB 271 GAGTGTGTAGCGGTGAGAGAAGATTTCCCTGACGAAGAAGATTTGTGAAAGCTGGTG 330
QY 74 ySerGluLeuLeuPheValGlnMetGlnThrLysSerMetGluLysGlnAlaLysLe 94
DB 331 TTCTGAGATTCTATTGTTTCAAAATGCAGCAACAAGATATGGATGACACTCTAAGCT 390
QY 94 uAlaAspLysLeuProPheIleProPheGlyGluSerValMetAspLeuValIleG1 114
DB 391 TGTGTATAGTTGCTCTATATCAATTTGGTGGTGGTCTAGTGGTATTGG 450
QY 114 yCysGlyProAlaGlyLeuSerLeuAlaAlaGluAlaLysLeuGlyLeuLysValG1 134
DB 451 TTGTGTCCTGCTGCTTTAGCTTGGCTGCAGATCAGCTAGCTGGATTGAAGATTGG 510
QY 134 yLeuIleGlyProAspLeuProPheThrAsnAsnTyrGlyValTrpGluAspGluPheLy 154
DB 511 ACTCATTTGCTCCAGATCTCTCTTTTACTTAACAATACGGTGTGGGAAGATGAATTCAA 570
QY 154 sAspLeuGlyLeuGluArgCysIleGluHisAlaTrpLysAspThrIleValTyrLeuAs 174
DB 154 sAspLeuGlyLeuGluArgCysIleGluHisAlaTrpLysAspThrIleValTyrLeuAs 174

```



```

Db 571 TGATCTGGGTGCAAAATGATTGACGATGTTTGGAGAGAGACTATTGNGTATCTGGA 630
Qy 174 pAsnAspAlaProValLeuIleGlyArgAlaTyrGly 186
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 631 TGATGACAGCCTATTACCATGTCGCGTGCTATGGA 667

RESULT 15
BJ482945/c
LOCUS
DEFINITION BJ482945 K. Sato unpublished cDNA library, strain H602 adult,
             heading stage top three leaves Hordeum vulgare subsp. spontaneum
             cDNA clone ban42d05 3', mRNA sequence.
ACCESSION
VERSION BJ482945.1 GI:21161400
KEYWORDS EST.
SOURCE
ORGANISM Hordeum vulgare subsp. spontaneum.
            Hordeum vulgare subsp. spontaneum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Hordeum.
REFERENCE 1 (bases 1 to 617)
AUTHORS Sato,K., Saisho,D. and Takeda,K.
TITLE Barley EST sequencing project in NIG and Okayama Univ
JOURNAL Unpublished (2002)
COMMENT Contact: Tadasu Shin-i
          Center For Genetic Resource Information
          National Institute of Genetics
          1111 Yata, Mishima, Shizuoka 411-8540, Japan
          Tel: 81-559-81-6856
          Fax: 81-559-81-6855
          Email: tshini@genes.nig.ac.jp.
FEATURES
            source
              1..617
                /organism="Hordeum vulgare subsp. spontaneum"
                /strain="H602"
                /db_xref="taxon:77009"
                /clone="ban42d05"
                /clone_lib="K. Sato unpublished cDNA library, strain H602
                adult, heading stage top three leaves"
                /tissue_type="top three leaves"
                /dev_stage="adult, heading stage"
BASE COUNT 181 a 137 c 143 g 155 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 67e-64 Length: 617
Score: 581.50 Matches: 112
Percent Similarity: 83.23% Conservative: 22
Best Local Similarity: 69.57% Mismatches: 26
Query Match: 21.20% Indels: 1
DB: 13 Gaps: 1

US-09-701-395A-23 (1-529) x BJ482945 (1-617)
Qy 368 AsnThrGluGlnLysAsnLeuAlaPheGlyAlaAalaSerMetValHisProAlaThr 387
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 615 AACACAGACCAGAAAATCTTGTCATTGGTGTGCGCAGCAGATGTCACCCCTGCAACT 556

Qy 388 GlyTyrSerValValArgSerLeuSerGluAlaProLysTyrAlaSerValIleAlaLys 407
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 555 GGATACTCGGTGGTCAGATCTTGTCTGAAGCTCCAGATATGCTCTGTGATATCTGAT 496

Qy 408 IleLeu---LysGlnAspAsnSerAlaTyrValValSerGlyGlnSerSerAlaValAsn 426
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 495 ATCTTAGGAATCGTCTATTCTGGACAATATTGCTGGAAGTTCGAAATGTCACGT 436

Qy 427 IleSerMetGlnAlaThrSerSerLeuTrpProLysGluArgLysArgGlnArgAlaPhe 446
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 435 CCATCAATGCTGCATGGGAAACACTATGGCCTCAAGAACGGAACGTCACGCTCATTC 376

Qy 447 PheLeuPheGlyLeuGluLeuIleValGlnLeuAspIleGluAlaThrArgThrPhePhe 466
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 375 TTCCTCTTTGGATTGGCCTTGATATTCACTGGAATACGAAGGATCCAAACGTTCTTC 316

```

```

Qy 467 ArgThrPheArgLeuProThrTyrMetTyrTyrPheLeuGlySerSerLeuSer 486
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 315 GAAACCTTTTCCGGTTACCAAAATGGATGTGGCGAGGATTCCTTGTGTCAGCGTGTG 256

Qy 487 SerPheAspLeuValLeuPheSerMetTyrMetPheValLeuAlaProAsnSerMetArg 506
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 255 TCAGTGGATCTCATGCTGTGTGCACTCTACATGTTTGCATTTGCAATTCGCGCAACAAATTTGCGA 196

Qy 507 MetSerLeuValArgHisLeuSerAspProSerGlyAlaValMetValArgAlaTyr 526
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 195 ATGAACCTCGTCAGACACCTCTCTCGGACCCCACTGTTTCAGCAATGATCAGGACCTAC 136

Qy 527 Leu 527
   |||
Db 135 CTG 133

```

Search completed: May 22, 2003, 00:16:00
Job time : 2156 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 21, 2003, 21:09:13 ; Search time 65 Seconds
(without alignments)
1084.455 Million cell updates/sec

Title: US-09-701-395A-23

Perfect score: 2743

Sequence: 1 MELLGVRNLISCFVWTEGT.....VRHLLSDPSGAVMVRAYLER 529

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :	A_Geneseq.101002.*
1:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.*
2:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.*
3:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.*
4:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.*
5:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.*
6:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.*
7:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT.*
8:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.*
9:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.*
10:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.*
11:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.*
12:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.*
13:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.*
14:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.*
15:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.*
16:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.*
17:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.*
18:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.*
19:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.*
20:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.*
21:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.*
22:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.*
23:	/SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2743	100.0	529	21	AA54290 An Adonis palaeostina lycopene epsilon-cyclase.
2	2743	100.0	529	21	AA54303 Amino acid sequence
3	2743	100.0	529	21	AA54306 Amino acid sequence
4	2728	99.5	529	21	AA54307 Amino acid sequence
5	2728	99.5	529	21	AA54309 Amino acid sequence
6	1966.5	71.7	524	23	AB93848 Herbicidally active
7	1959.5	71.4	524	18	AAW35346 Arabidopsis thaliana
8	1959.5	71.4	524	21	AAW54283 Amino acid sequence
9	1959.5	71.4	524	21	AA54305 Amino acid sequence
10	1934	70.5	516	21	AA54290 Arabidopsis thaliana

11	1934	70.5	516	21	AA54310 Amino acid sequence
12	1934	70.5	533	21	AA54308 Amino acid sequence
13	1925.5	70.2	532	21	AA52592 Romaine lettuce ep
14	1908.5	69.6	526	21	AA54309 Amino acid sequence
15	1892	69.0	517	22	AAE09798 Spinach lycopene e
16	1882	68.6	533	21	AA54292 Sequence of a chim
17	1518.5	55.4	378	21	AA54291 A potato lycopene
18	1518.5	55.4	378	21	AA54304 Amino acid sequence
19	1166	42.5	317	21	AA53388 Soybean lycopene e
20	1157.5	42.2	337	21	AA53388 Arabidopsis thaliana
21	1157.5	42.2	345	21	AA53388 Arabidopsis thaliana
22	1142.5	41.7	326	21	AA53388 Arabidopsis thaliana
23	775.5	28.3	502	21	AA54312 Amino acid sequence
24	761	27.7	498	18	AAW06454 Capsicum annuum ly
25	761	27.7	498	21	AA54313 Amino acid sequence
26	755	27.5	511	21	AA54316 Amino acid sequence
27	750	27.3	511	21	AA54316 Arabidopsis thaliana
28	749.5	27.3	500	21	AA54314 Arabidopsis thaliana
29	748	27.3	501	21	AA54311 Arabidopsis thaliana
30	748	27.3	501	21	AA54311 Arabidopsis thaliana
31	748	27.3	501	23	AB93263 Herbicidally active
32	733	26.7	500	21	AA54315 Amino acid sequence
33	722.5	26.3	498	21	AA54315 Arabidopsis thaliana
34	719.5	26.2	498	21	AA54315 Arabidopsis thaliana
35	719.5	26.2	498	21	AA54315 Arabidopsis thaliana
36	705.5	25.7	498	16	AA82242 Capsanthin-capsoru
37	701.5	25.6	456	21	AA54289 Consensus sequence
38	701.5	25.6	503	21	AA54317 Amino acid sequence
39	637.5	23.2	493	21	AA53325 Soybean lycopene c
40	620.5	22.6	369	21	AA53325 Arabidopsis thaliana
41	605	22.1	411	16	AAW0276 Lycopene cyclase.
42	605	22.1	411	17	AAW0276 Cyanobacterial lyc
43	585.5	21.3	185	21	AA53387 Corn lycopene epsi
44	562.5	20.5	353	21	AA53324 Corn lycopene cycl
45	551.5	20.1	328	21	AA54313 Arabidopsis thaliana

ALIGNMENTS

RESULT 1	AA54290
ID	AA54290 standard; Protein; 529 AA.
XX	
AC	AA54290;
XX	
DT	06-APR-2000 (first entry)
XX	
DE	An Adonis palaeostina lycopene epsilon-cyclase.
XX	
KW	Lycopene epsilon-cyclase; carotenoid biosynthesis pathway;
KW	isopentenyl pyrophosphate isomerase; beta-carotene hydroxylase;
KW	dimethylallyl pyrophosphate; transgenic plant; carotenoid content;
KW	lethal photooxidative damage; colouring.
XX	
OS	Adonis palaeostina.
XX	
PN	WO9963055-A1.
XX	
PD	09-DEC-1999.
XX	
PF	02-JUN-1999; 99WO-US12121.
XX	
PR	02-JUN-1998; 98US-0088724.
XX	
PA	02-JUN-1998; 98US-0088725.
XX	
PI	(UYMA-) UNIV MARYLAND BALTIMORE.
XX	
DR	Cunningham FX, Sun Z;
XX	
DR	WPI; 2000-116366/10.
XX	
DR	N-PSDB; AA245585.
XX	

PT New carotenoid biosynthesis genes used to provide transgenic plants
 PT with altered carotenoid content
 XX Claim 2; Fig 14; 134pp; English.
 PS
 CC The present sequence represents a lycopene epsilon-cyclase protein.
 CC The enzyme is involved in the carotenoid biosynthesis pathway. The
 CC specification also describes other enzymes involved in this pathway,
 CC such as epsilon-cyclase, isopentenyl pyrophosphate (IPP) isomerase and
 CC beta-carotene hydroxylase. IPP isomerase catalyses the reversible
 CC conversion of IPP to dimethylallyl pyrophosphate (DMAPP). Beta-carotene
 CC hydroxylase is responsible for hydroxylating the beta-endgroup in
 CC carotenoids. Lycopene epsilon-cyclase is responsible for the formation
 CC of epsilon-endgroups in carotenoids. The polynucleotides of the
 CC invention are used to provide transgenic plants with altered
 CC carotenoid content. These plants may, for example, may be protected
 CC from lethal photooxidative damage or have altered colouring.
 XX
 SQ Sequence 529 AA;
 Query Match 100.0%; Score 2743; DB 21; Length 529;
 Best Local Similarity 100.0%; Pred. No. 1.4e-271;
 Matches 529; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MELLGVRNLISCPVWTFGTNRNLSSSKLAYNIHRYGSSCRVDFQVRADGGSSRSVAYK 60
 DB 1 MELLGVRNLISCPVWTFGTNRNLSSSKLAYNIHRYGSSCRVDFQVRADGGSSRSVAYK 60
 QY 61 EGFVDEEDFKAGSELFLVQMOTKSMKQAKLADKLPIPFGESVMDLVWIGCGPAGL 120
 DB 61 EGFVDEEDFKAGSELFLVQMOTKSMKQAKLADKLPIPFGESVMDLVWIGCGPAGL 120
 QY 121 SLAAEAAKGLKVLGDLPGDLPFTNNYGVWEDEFKDLGLERCIEHAWKDTIVYLDNDAPVL 180
 DB 121 SLAAEAAKGLKVLGDLPGDLPFTNNYGVWEDEFKDLGLERCIEHAWKDTIVYLDNDAPVL 180
 QY 181 IGRAYGRVSRHLLHEELLKRCVSGSYLDSKVERITEAGDGHSLVVCENEIIFPCRLAT 240
 DB 181 IGRAYGRVSRHLLHEELLKRCVSGSYLDSKVERITEAGDGHSLVVCENEIIFPCRLAT 240
 QY 241 VASGAASGKLLYEYGVGPRVCQTAYGVEVEENPDNLMVPMYRDYMQKLQCSSE 300
 DB 241 VASGAASGKLLYEYGVGPRVCQTAYGVEVEENPDNLMVPMYRDYMQKLQCSSE 300
 QY 301 EYPTFLYMPSPRLFFETCTCLASKDAMPFDLLKRLMSRLKTLGIQVTKVYEEESYI 360
 DB 301 EYPTFLYMPSPRLFFETCTCLASKDAMPFDLLKRLMSRLKTLGIQVTKVYEEESYI 360
 QY 361 PVGSLPNTQKNLAFGAASMVHPATGYSVYRSLSEAPKYASVIAKILKQDINSAYVVG 420
 DB 361 PVGSLPNTQKNLAFGAASMVHPATGYSVYRSLSEAPKYASVIAKILKQDINSAYVVG 420
 QY 421 QSSAVNTSMOAWSLWPKERRKORAFFLGLGLIVQLDIEATRTFFTFRLPTWMMWGF 480
 DB 421 QSSAVNTSMOAWSLWPKERRKORAFFLGLGLIVQLDIEATRTFFTFRLPTWMMWGF 480
 QY 481 LGSSLSFDLVLFNSYMFVLAPNSMRSLVRHLLSDPSGAVMVRAYLER 529
 DB 481 LGSSLSFDLVLFNSYMFVLAPNSMRSLVRHLLSDPSGAVMVRAYLER 529
 RESULT 2
 AAY54303
 ID AAY54303 standard; Protein; 529 AA.
 XX
 AC AAY54303;
 CC
 DT 06-APR-2000 (first entry)
 XX
 DE An Adonis palaeatina lycopene epsilon-cyclase.
 XX Lycopene epsilon-cyclase; carotenoid biosynthesis pathway;
 KW isopentenyl pyrophosphate isomerase; beta-carotene hydroxylase;

KW dimethylallyl pyrophosphate; transgenic plant; carotenoid content;
 KW lethal photooxidative damage; colouring.
 XX Adonis palaeatina.
 OS
 XX
 PH Key Location/Qualifiers
 FT Misc-difference 55 /note= "encoded by ACT"
 FT Misc-difference 210 /note= "encoded by AAT"
 FT Misc-difference 231 /note= "encoded by GAC"
 FT Misc-difference 352 /note= "encoded by ATT"
 FT Misc-difference 524 /note= "encoded by AAA"
 XX
 PN W09963055-A1.
 XX
 PD 09-DEC-1999.
 XX
 PF 02-JUN-1999; 99WO-US12121.
 XX
 PR 02-JUN-1998; 98US-0088724.
 PR 02-JUN-1998; 98US-0088725.
 XX
 PA (UYMA-) UNIV MARYLAND BALTIMORE.
 XX
 PI Cunningham FX, Sun Z;
 XX WPI; 2000-116366/10.
 DR N-PSDB; AA245595.
 DR
 XX New carotenoid biosynthesis genes used to provide transgenic plants
 with altered carotenoid content
 PS Disclosure; Fig 24; 134pp; English.
 XX
 CC The present sequence represents a lycopene epsilon-cyclase protein.
 CC The enzyme is involved in the carotenoid biosynthesis pathway. The
 CC specification also describes other enzymes involved in this pathway,
 CC such as epsilon-cyclase, isopentenyl pyrophosphate (IPP) isomerase and
 CC beta-carotene hydroxylase. IPP isomerase catalyses the reversible
 CC conversion of IPP to dimethylallyl pyrophosphate (DMAPP). Beta-carotene
 CC hydroxylase is responsible for hydroxylating the beta-endgroup in
 CC carotenoids. Lycopene epsilon-cyclase is responsible for the formation
 CC of epsilon-endgroups in carotenoids. The polynucleotides of the
 CC invention are used to provide transgenic plants with altered
 CC carotenoid content. These plants may, for example, may be protected
 CC from lethal photooxidative damage or have altered colouring.
 XX
 SQ Sequence 529 AA;
 Query Match 100.0%; Score 2743; DB 21; Length 529;
 Best Local Similarity 100.0%; Pred. No. 1.4e-271;
 Matches 529; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MELLGVRNLISCPVWTFGTNRNLSSSKLAYNIHRYGSSCRVDFQVRADGGSSRSVAYK 60
 DB 1 MELLGVRNLISCPVWTFGTNRNLSSSKLAYNIHRYGSSCRVDFQVRADGGSSRSVAYK 60
 QY 61 EGFVDEEDFKAGSELFLVQMOTKSMKQAKLADKLPIPFGESVMDLVWIGCGPAGL 120
 DB 61 EGFVDEEDFKAGSELFLVQMOTKSMKQAKLADKLPIPFGESVMDLVWIGCGPAGL 120
 QY 121 SLAAEAAKGLKVLGDLPGDLPFTNNYGVWEDEFKDLGLERCIEHAWKDTIVYLDNDAPVL 180
 DB 121 SLAAEAAKGLKVLGDLPGDLPFTNNYGVWEDEFKDLGLERCIEHAWKDTIVYLDNDAPVL 180
 QY 181 IGRAYGRVSRHLLHEELLKRCVSGSYLDSKVERITEAGDGHSLVVCENEIIFPCRLAT 240
 DB 181 IGRAYGRVSRHLLHEELLKRCVSGSYLDSKVERITEAGDGHSLVVCENEIIFPCRLAT 240

QY 241 VASGAASGKLLLEYEGGPRVCVQTAYGVEVEVENNPDNLMVMDYRDYMOOKLQCSSE 300
 Db 241 VASGAASGKLLLEYEGGPRVCVQTAYGVEVEVENNPDNLMVMDYRDYMOOKLQCSSE 300
 QY 301 EYPTFLYVMPSPTRLFFETCTCLASKDAMPFDLLKRLKSLRLKTLGIGQVTKYEEWSYI 360
 Db 301 EYPTFLYVMPSPTRLFFETCTCLASKDAMPFDLLKRLKSLRLKTLGIGQVTKYEEWSYI 360
 QY 361 PYGGSPLNTEQKNLAFGAASVHPATGYSVVRSLSEAPKYASVIAKILKQNSAYVWSG 420
 Db 361 PYGGSPLNTEQKNLAFGAASVHPATGYSVVRSLSEAPKYASVIAKILKQNSAYVWSG 420
 QY 421 QSSAVNISMQANSSLPKPKRQRAFFLGLLELIVOLDIEATRTFFRPLPTMMWGF 480
 Db 421 QSSAVNISMQANSSLPKPKRQRAFFLGLLELIVOLDIEATRTFFRPLPTMMWGF 480
 QY 481 LGSSLSFDLVLFMSYMFVLAPNSMRMSLVRHLLSDPSGAVMVAYLER 529
 Db 481 LGSSLSFDLVLFMSYMFVLAPNSMRMSLVRHLLSDPSGAVMVAYLER 529
 RESULT 3
 AAY54306
 ID AAY54306 standard; Protein; 529 AA.
 XX
 AC AAY54306;
 XX
 DT 06-APR-2000 (first entry)
 XX
 DE Amino acid sequence of a lycopene epsilon-cyclase protein.
 XX
 KW Lycopene epsilon-cyclase; carotenoid biosynthesis pathway;
 KW isopentenyl pyrophosphate isomerase; beta-carotene hydroxylase;
 KW dimethylallyl pyrophosphate; transgenic plant; carotenoid content;
 KW lethal photooxidative damage; colouring.
 XX
 OS Adonis palaestina.
 XX
 PN WO9963055-A1.
 XX
 PD 09-DEC-1999.
 XX
 PF 02-JUN-1999; 99WO-US12121.
 XX
 PR 02-JUN-1998; 98US-0088724.
 PR 02-JUN-1998; 98US-0088725.
 XX
 PA (UYMA-) UNIV MARYLAND BALTIMORE.
 XX
 PI Cunningham FX, Sun Z;
 XX
 DR WPI; 2000-116366/10.
 XX
 PT New carotenoid biosynthesis genes used to provide transgenic plants
 with altered carotenoid content -
 XX
 PS Disclosure; Fig 25; 134pp; English.
 XX
 CC The present sequence represents a lycopene epsilon-cyclase protein.
 CC The enzyme is involved in the carotenoid biosynthesis pathway. The
 CC specification also describes other enzymes involved in this pathway,
 CC such as epsilon-cyclase, isopentenyl pyrophosphate (IPP) isomerase and
 CC beta-carotene hydroxylase. IPP isomerase catalyses the reversible
 CC conversion of IPP to dimethylallyl pyrophosphate (DMAPP). Beta-carotene
 CC hydroxylase is responsible for hydroxylating the beta-endgroup in
 CC carotenoids. Lycopene epsilon-cyclase is responsible for the formation
 CC of epsilon-endgroups in carotenoids. The polynucleotides of the
 CC invention are used to provide transgenic plants with altered
 CC carotenoid content. These plants may, for example, may be protected
 CC from lethal photooxidative damage or have altered colouring.
 XX
 SQ Sequence 529 AA;

Query Match 100.0%; Score 2743; DB 21; Length 529;
 Best Local Similarity 100.0%; Pred. No. 1.4e-271;
 Matches 529; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MELLGVRNLISSCPVWTFGRNLSSSKLAYNTHRYGSSCRVDQVQVADGGSGRSRVAYK 60
 Db 1 MELLGVRNLISSCPVWTFGRNLSSSKLAYNTHRYGSSCRVDQVQVADGGSGRSRVAYK 60
 QY 61 EGFVDEEDFIKAGGSLLFVQMOOTKSMKQAKLADKLPPIPGESVMDLVVIGCCPAGL 120
 Db 61 EGFVDEEDFIKAGGSLLFVQMOOTKSMKQAKLADKLPPIPGESVMDLVVIGCCPAGL 120
 QY 121 SLAAEAAKLGVLGIGLDPDLPFTNNYGVWEDEFKDLGLERCIEHAWKDTIVYLDNDAPVL 180
 Db 121 SLAAEAAKLGVLGIGLDPDLPFTNNYGVWEDEFKDLGLERCIEHAWKDTIVYLDNDAPVL 180
 QY 181 IGRAYGRVSRHLLHEELLKRCVSGVSYLDSKVERITEAGDGHSLVVCENEIPICRLAT 240
 Db 181 IGRAYGRVSRHLLHEELLKRCVSGVSYLDSKVERITEAGDGHSLVVCENEIPICRLAT 240
 QY 241 VASGAASGKLLLEYEGGPRVCVQTAYGVEVEVENNPDNLMVMDYRDYMOOKLQCSSE 300
 Db 241 VASGAASGKLLLEYEGGPRVCVQTAYGVEVEVENNPDNLMVMDYRDYMOOKLQCSSE 300
 QY 301 EYPTFLYVMPSPTRLFFETCTCLASKDAMPFDLLKRLKSLRLKTLGIGQVTKYEEWSYI 360
 Db 301 EYPTFLYVMPSPTRLFFETCTCLASKDAMPFDLLKRLKSLRLKTLGIGQVTKYEEWSYI 360
 QY 361 PYGGSPLNTEQKNLAFGAASVHPATGYSVVRSLSEAPKYASVIAKILKQNSAYVWSG 420
 Db 361 PYGGSPLNTEQKNLAFGAASVHPATGYSVVRSLSEAPKYASVIAKILKQNSAYVWSG 420
 QY 421 QSSAVNISMQANSSLPKPKRQRAFFLGLLELIVOLDIEATRTFFRPLPTMMWGF 480
 Db 421 QSSAVNISMQANSSLPKPKRQRAFFLGLLELIVOLDIEATRTFFRPLPTMMWGF 480
 QY 481 LGSSLSFDLVLFMSYMFVLAPNSMRMSLVRHLLSDPSGAVMVAYLER 529
 Db 481 LGSSLSFDLVLFMSYMFVLAPNSMRMSLVRHLLSDPSGAVMVAYLER 529
 RESULT 4
 AAY54307
 ID AAY54307 standard; Protein; 529 AA.
 XX
 AC AAY54307;
 XX
 DT 06-APR-2000 (first entry)
 XX
 DE Amino acid sequence of a lycopene epsilon-cyclase protein.
 XX
 KW Lycopene epsilon-cyclase; carotenoid biosynthesis pathway;
 KW isopentenyl pyrophosphate isomerase; beta-carotene hydroxylase;
 KW dimethylallyl pyrophosphate; transgenic plant; carotenoid content;
 KW lethal photooxidative damage; colouring.
 XX
 OS Adonis palaestina.
 XX
 PN WO9963055-A1.
 XX
 PD 09-DEC-1999.
 XX
 PF 02-JUN-1999; 99WO-US12121.
 XX
 PR 02-JUN-1998; 98US-0088724.
 PR 02-JUN-1998; 98US-0088725.
 XX
 PA (UYMA-) UNIV MARYLAND BALTIMORE.
 XX
 PI Cunningham FX, Sun Z;
 XX
 DR WPI; 2000-116366/10.
 XX

PT New carotenoid biosynthesis genes used to provide transgenic plants
 XX with altered carotenoid content
 PS Disclosure; Fig 25; 134pp; English.
 XX
 CC The present sequence represents a lycopene epsilon-cyclase protein.
 CC The enzyme is involved in the carotenoid biosynthesis pathway. The
 CC specification also describes other enzymes involved in this pathway,
 CC such as epsilon-cyclase, isopentenyl pyrophosphate (IPP) isomerase and
 CC beta-carotene hydroxylase. IPP isomerase catalyses the reversible
 CC conversion of IPP to dimethylallyl pyrophosphate (DMAPP). Beta-carotene
 CC hydroxylase is responsible for hydroxylating the beta-endgroup in
 CC carotenoids. Lycopene epsilon-cyclase is responsible for the formation
 CC of epsilon-endgroups in carotenoids. The polynucleotides of the
 CC invention are used to provide transgenic plants with altered
 CC carotenoid content. These plants may, for example, may be protected
 CC from lethal photooxidative damage or have altered colouring.
 XX
 SQ Sequence 529 AA;
 Query Match 99.5%; Score 2728; DB 21; Length 529;
 Best Local Similarity 99.1%; Pred. No. 4.9e-270;
 Matches 524; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MELLGVRNLISCPWTFGTNRNLSSKLAYNIHRYGSSCRVDFQVRADGGSGSRSSVAYK 60
 DB 1 MELLGVRNLISCPWTFGTNRNLSSKLAYNIHRYGSSCRVDFQVRADGGSGSRSSVAYK 60
 QY 61 EGFVDEEDFIKAGSELFLVQMQOTKMEKQAKLADKLPIPPFGESVMDLVVIGCGPAGL 120
 DB 61 EGFVDEEDFIKAGSELFLVQMQOTKMEKQAKLADKLPIPPFGESVMDLVVIGCGPAGL 120
 QY 121 SLAAEAAKGLKVLGIPDLPTFTNNYGVWEDEFKDLGLERIEHAWKDTIVYLDNDAPVL 180
 DB 121 SLAAEAAKGLKVLGIPDLPTFTNNYGVWEDEFKDLGLERIEHAWKDTIVYLDNDAPVL 180
 QY 181 IGRAYGRVSRHLLHEELLKRCVSGSVYLSKVERITEAGDGHSLVVCENEIFPCRLAT 240
 DB 181 IGRAYGRVSRHLLHEELLKRCVSGSVYLSKVERITEAGDGHSLVVCENEIFPCRLAT 240
 QY 241 VASGAASGKLLLEYVGGPRVCQQTAYGVEVEVENNPDNLMFMDYRDYMQQKLCQSEE 300
 DB 241 VASGAASGKLLLEYVGGPRVCQQTAYGVEVEVENNPDNLMFMDYRDYMQQKLCQSEE 300
 QY 301 EYPTFLYMPSPRLFEETCTCLASKDAMPDILLKRLMSRLKTLGIQVTKIYEENSYI 360
 DB 301 EYPTFLYMPSPRLFEETCTCLASKDAMPDILLKRLMSRLKTLGIQVTKIYEENSYI 360
 QY 361 PVGSLPNTQKNLAFGAASMVHPATGYSVRSLSSEAPKVASVIKILKODNSAYVYVG 420
 DB 361 PVGSLPNTQKNLAFGAASMVHPATGYSVRSLSSEAPKVASVIKILKODNSAYVYVG 420
 QY 421 QSSAVNISMQAWSLWPKERRKQRAFFLGLLELVQIDIEATRTFFTRFPLPTMMWGF 480
 DB 421 QSSAVNISMQAWSLWPKERRKQRAFFLGLLELVQIDIEATRTFFTRFPLPTMMWGF 480
 QY 481 LGSLSLSDFLVFSMYMFVLAPNSMRSLVRLHLLSDPSGAVMVYAYLER 529
 DB 481 LGSLSLSDFLVFSMYMFVLAPNSMRSLVRLHLLSDPSGAVMVYAYLER 529
 RESULT 5
 AAY52593
 ID AAY52593 standard; Protein: 529 AA.
 XX
 AC AAY52593;
 CC
 DT 07-MAR-2000 (first entry)
 XX
 DE Adonis palaestina epsilon lycopene cyclase.
 XX Lycopene; epsilon cyclase; bicyclic epsilon carotene; biosynthesis;
 KW carotenoid; epsilon ring; lactucaxanthin; protein production;
 KW

KW transgenic plant; altered synthesis; overexpression;
 KW nutritional value; pharmacology; colour; cancer; antisense;
 KW accumulation; substrate; beta-carotene; vitamin A.
 XX
 OS Adonis palaestina.
 XX
 PN WO9961399-A1.
 XX
 PD 02-DEC-1999.
 XX
 PF 25-MAY-1999; 99WO-US10461.
 XX
 PR 26-MAY-1998; 98US-0084222.
 XX
 PA (UYMA-) UNIV MARYLAND BALTIMORE.
 XX
 PI Cunningham FX;
 XX
 DR WPI; 2000-062667/05.
 DR N-PSDB; AAZ46329.
 XX
 PT New eukaryotic epsilon-lycopene cyclase and related DNA, used to
 PT regulate carotenoid synthesis in plants and other hosts, e.g. for
 PT treatment of cancer
 XX
 PS Disclosure; Fig 6B; 38pp; English.
 XX
 CC This sequence represents epsilon lycopene cyclase from Adonis
 CC palaestina. This enzyme adds two epsilon rings to lycopene to
 CC form bicyclic epsilon-carotene. Carotenoids with two epsilon rings
 CC are found in significant amounts in relatively few plants (romaine
 CC lettuce produces an abundance of such a carotenoid, lactucaxanthin),
 CC while carotenoids with two beta rings are ubiquitous and those
 CC with one beta and one epsilon ring are common. The invention relates to
 CC epsilon lycopene cyclases capable of forming bicyclic epsilon-carotene,
 CC and also to methods of transforming known carotenoids into novel or rare
 CC products. Epsilon lycopene cyclase nucleic acids may be used for
 CC recombinant protein production and for the generation of transgenic
 CC plants that have altered patterns of carotenoid synthesis. Such plants
 CC may accumulate new or rare carotenoids or overexpress known carotenoids,
 CC resulting in altered nutritional value, pharmacology or colour.
 CC Carotenoids isolated from such transgenic plants are useful
 CC therapeutically, e.g., for treatment of cancer. Alternatively, epsilon
 CC lycopene cyclase antisense nucleic acids can be used to cause
 CC accumulation of the substrates of the enzyme, specifically beta-carotene
 CC or derived products such as vitamin A. The isolated eukaryotic enzyme
 CC allows synthesis of epsilon, epsilon-carotene (which can be
 CC isolated from natural sources only in small amounts) on a large scale,
 CC enabling bulk production of derived carotenoids.
 XX
 SQ Sequence 529 AA;
 Query Match 99.5%; Score 2728; DB 21; Length 529;
 Best Local Similarity 99.1%; Pred. No. 4.9e-270;
 Matches 524; Conservative 5; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MELLGVRNLISCPWTFGTNRNLSSKLAYNIHRYGSSCRVDFQVRADGGSGSRSSVAYK 60
 DB 1 MELLGVRNLISCPWTFGTNRNLSSKLAYNIHRYGSSCRVDFQVRADGGSGSRSSVAYK 60
 QY 61 EGFVDEEDFIKAGSELFLVQMQOTKMEKQAKLADKLPIPPFGESVMDLVVIGCGPAGL 120
 DB 61 EGFVDEEDFIKAGSELFLVQMQOTKMEKQAKLADKLPIPPFGESVMDLVVIGCGPAGL 120
 QY 121 SLAAEAAKGLKVLGIPDLPTFTNNYGVWEDEFKDLGLERIEHAWKDTIVYLDNDAPVL 180
 DB 121 SLAAEAAKGLKVLGIPDLPTFTNNYGVWEDEFKDLGLERIEHAWKDTIVYLDNDAPVL 180
 QY 181 IGRAYGRVSRHLLHEELLKRCVSGSVYLSKVERITEAGDGHSLVVCENEIFPCRLAT 240
 DB 181 IGRAYGRVSRHLLHEELLKRCVSGSVYLSKVERITEAGDGHSLVVCENEIFPCRLAT 240
 QY 241 VASGAASGKLLLEYVGGPRVCQQTAYGVEVEVENNPDNLMFMDYRDYMQQKLCQSEE 300

Db	1	MECVGARNAAMAVSTFPNSCSRKKFPVVKRYSYRNIRFG-LC-----SVRASGGSGSGSE	234
QY	55	SSVAYKEGFVDEEDFKAGSGELLFVOMQOTKSMKQAKLADKLPPITPFGESVMDLWVIG	114
Db	56	SCVAVREDFADEDFYKAGSEILFVOMQONKMDQSKLYDKLPPISIGDGLDVLWIG	115
QY	115	CGPAGLSAABAAKGLKVLGIGDLPFTNNYGWGEDFKDLGLERCIEHAWKDTIVYLD	174
Db	116	CGPAGLALAAESAKGLKVLGIGDLPFTNNYGWGEDFNLDGLKCIHVVWRETIVYLD	175
QY	175	NDAPVLIGRAYGRVSRHLLHEELLKRCVESGVSLDSKVERITEAGDGHSLVVCENEIFI	234
Db	176	DDPITIGRAYGRVSRHLLHEELLKRCVESGVSLDSKVERITEAGDGLRVACDDNNVI	235
QY	235	PCRLATVASGAASKLLEYEVGGPRVCVQTAYGVEVEVENNPYDPNLMVFM DYRDMQOK	294
Db	236	PCRLATVASGAASKLLQYEVGGPRVCVQTAYGVEVEVENSPYDPQVFM DYRYTNEK	295
QY	295	LQCSSEYPTFLVMPMSPTPLFEETCLASKAMPDDLKRLKMSRLKTLGLQVTKVYE	354
Db	296	VRSLAEYPTFLYAMPMTKSRLEFEETCLASKQVMPDDLKTLMLRLDLTLGRILKTYE	355
QY	355	EESYIPVGGSLNTEOKNLAFGAASAMVHPATGYSVVRSLSEAPKVASVIAKILKODNS	414
Db	356	EESYIPVGGSLNTEOKNLAFGAASAMVHPATGYSVVRSLSEAPKVASVIAELREETT	415
QY	415	AYVVGSSAVNISMOAWSLWPKERRQRRAFFLFGLELIVQLDIEATRTFFTRFLPT	474
Db	416	KQNS-----NLSRQAWDTLWPKERRQRRAFFLFGLALIVQDTGIRSFFFTRFLPK	469
QY	475	WMWGFGLGSLSDFLVLFWSYMFVLA PMSRMSLVRLHLLSDPSGAVMVRAYLE	528
Db	476	WMWGFGLGSLSDFLVLFYALYMFVISPNNLRKGLINHLISDPTGATMIKTYLK	523
RESULT	7		
AAW35346			
ID	AAW35346	standard; Protein; 524 AA.	
XX	AAW35346;		
AC			
XX			
DT	14-APR-1998	(first entry)	
XX			
DE	Arabidopsis thaliana epsilon cyclase.		
DE			
KW	Epsilon cyclase; E-cyclase; carotenoid; biosynthetic enzyme;		
KW	pigment; vector; pAtEps.		
XX			
OS	Arabidopsis thaliana.		
XX			
PN	WO9736998-A1.		
XX			
PD	09-OCT-1997.		
XX			
PF	28-JAN-1997; 97WO-US00540.		
XX			
PR	29-MAR-1996; 96US-0624125.		
XX			
PA	(UYMA-) UNIV MARYLAND BALTIMORE.		
XX			
PI	Cunningham FX, Sun Z;		
XX			
DR	WPI; 1997-503091/46.		
DR	N-PSDB; AAT95371.		
XX			
PT	Eukaryotic carotenoid biosynthetic enzymes and related genes -		
PT	useful to control ratio of various carotenoid(s) in host and for		
PT	production of novel carotenoid pigments		
XX			
PS	Claim 1; Page 33-35; 89pp; English.		
XX			
CC	This protein comprises Arabidopsis thaliana epsilon cyclase, an		
CC	enzyme responsible for the formation of epsilon end-groups in		

CC carotenoids. Its amino acid sequence was deduced from an
 CC isolated cDNA clone (see AAT95371). Expression vector pATEPS
 CC comprising the epsilon cyclase gene is deposited as ATCC 98005.
 CC The claimed eukaryotic carotenoid biosynthetic enzymes
 CC epsilon cyclase, beta-carotene hydroxylase and isopentenyl
 CC pyrophosphate isomerase (see AAW35346-51) are used in methods for
 CC augmenting the accumulation of carotenoids and for the production
 CC of novel and rare carotenoids in host cells. Methods are also
 CC provided for controlling the ratio of various carotenoids in a
 CC host, and for screening for eukaryotic genes that encode enzymes
 CC of carotenoid biosynthesis and metabolism.
 XX
 XX Sequence 524 AA;
 SQ
 Query Match 71.4%; Score 1959.5; DB 18; Length 524;
 Best Local Similarity 70.4%; Pred. No. 2.5e-191;
 Matches 376; Conservative 60; Mismatches 81; Indels 17; Gaps 5;
 QY 1 MELLGVRNL-----ISSCPWTFGTNRNLSSSLAYNIHRYGSSCRVDFQVRDGG--SGSR 54
 DB 1 MECVGARNAFAAMAVSTFSSWCRKRPVVKRYSYRNIIRFG-LC-----SVRASGGSGSGSE 55
 QY 55 SSVAKEGVDEEDFIKAGGSELLFVQMOQTKSMQAKLADKLPPIPFSGESVMDLVVIG 114
 DB 56 SCVAVREDAEEDFVRAAGSEILFVQMOQKMDQESKLVKLPPIGIGDGLDHHVIG 115
 QY 115 CGPAGLSAAEAAGLGLKVLGDPDPPTNNYGVWEDEFKDLGLERCIEHAWKDTIVYLD 174
 DB 116 CGPAGLAAEAAGLGLKVLGDPDPPTNNYGVWEDEFKDLGLERCIEHAWKDTIVYLD 175
 QY 175 NDAPVLIGRAYGRVSRHLLHEELLKRCVSGVSYLDSKVRITERAGDGHSLVVCENEIFI 234
 DB 176 DDKPITIGRAYGRVSRHLLHEELLKRCVSGVSYLDSKVRITERAGDGHSLVVCENEIFI 235
 QY 235 PCRLATVASGAASGLLEVEYGVGGPRVCVQTAYGVEVEVENNPDPNLMVMDYRDMOQK 294
 DB 236 PCRLATVASGAASGLLQYEVGGPRVCVQTAYGVEVEVENNPDPNLMVMDYRDMOQK 295
 QY 295 LQCSSEYPTFLYVMPKTRLFEEETCLASKDAMPDGLLRKMSRLTKLGIGQVTKYE 354
 DB 296 VRSLEAEYPTFLYVMPKTRLFEEETCLASKDAMPDGLLRKMSRLTKLGIGQVTKYE 355
 QY 355 EWSYIPVGGSLPNTQKNLAFGAASMVHPATGYSVVRSLSEAPKYASVIAKILKQDNS 414
 DB 356 EWSYIPVGGSLPNTQKNLAFGAASMVHPATGYSVVRSLSEAPKYASVIAKILRETT 415
 QY 415 AYVVSQSSAVNISQMAWSSLPKRRQRAFFLGLGLIIVQDIEATRTFFRFLPT 474
 DB 416 KQINS-----NISQAWDTLWPPERKQRAFFLGLGLIIVQDIEATRTFFRFLPK 469
 QY 475 WNWGFLGSSLSFGLVLFMYMFLAPNSMRSLVRLHLLSDPQGVAMVRAYLE 528
 DB 470 WNWQFLGSLTSGDLVLFALYMEVISPNNLRKGLNHLISDPTGATIKTYLK 523
 RESULT 8
 AAY54283
 ID AAY54283 standard; Protein; 524 AA.
 XX
 XX AAY54283;
 XX
 XX 06-APR-2000 (first entry)
 XX
 DE An Arabidopsis thaliana lycopene epsilon-cyclase.
 XX
 KW Lycopene epsilon-cyclase; carotenoid biosynthesis pathway;
 KW isopentenyl pyrophosphate isomerase; beta-carotene hydroxylase;
 KW dimethylallyl pyrophosphate; transgenic plant; carotenoid content;
 KW lethal photooxidative damage; colouring.
 XX
 OS Arabidopsis thaliana.
 XX
 PN W09963055-A1.

XX 09-DEC-1999.
 XX 02-JUN-1999; 99WO-US12121.
 XX 02-JUN-1998; 98US-0088724.
 XX 02-JUN-1998; 98US-0088725.
 PA (UYMA-) UNIV MARYLAND BALTIMORE.
 XX
 PI Cunningham FX, Sun Z;
 XX
 DR WPI; 2000-116366/10.
 DR N-PSDB; AAZ45578.
 XX
 PT New carotenoid biosynthesis genes used to provide transgenic plants
 with altered carotenoid content -
 XX
 PS Example; Page 64-65; 134pp; English.
 XX
 CC The present sequence represents a lycopene epsilon-cyclase protein.
 CC The enzyme is involved in the carotenoid biosynthesis pathway. The
 CC specification also describes other enzymes involved in this pathway,
 CC such as epsilon-cyclase, isopentenyl pyrophosphate (IPP) isomerase and
 CC beta-carotene hydroxylase. IPP isomerase catalyses the reversible
 CC conversion of IPP to dimethylallyl pyrophosphate (DMAPP). Beta-carotene
 CC hydroxylase is responsible for hydroxylating the beta-endgroup in
 CC carotenoids. Lycopene epsilon-cyclase is responsible for the formation
 CC of epsilon-endgroups in carotenoids. The polynucleotides of the
 CC invention are used to provide transgenic plants with altered
 CC carotenoid content. These plants may, for example, may be protected
 CC from lethal photooxidative damage or have altered colouring.
 XX
 SQ Sequence 524 AA;
 Query Match 71.4%; Score 1959.5; DB 21; Length 524;
 Best Local Similarity 70.4%; Pred. No. 2.5e-191;
 Matches 376; Conservative 60; Mismatches 81; Indels 17; Gaps 5;
 QY 1 MELLGVRNL-----ISSCPWTFGTNRNLSSSLAYNIHRYGSSCRVDFQVRDGG--SGSR 54
 DB 1 MECVGARNAFAAMAVSTFSSWCRKRPVVKRYSYRNIIRFG-LC-----SVRASGGSGSGSE 55
 QY 55 SSVAKEGVDEEDFIKAGGSELLFVQMOQTKSMQAKLADKLPPIPFSGESVMDLVVIG 114
 DB 56 SCVAVREDAEEDFVRAAGSEILFVQMOQKMDQESKLVKLPPIGIGDGLDHHVIG 115
 QY 115 CGPAGLSAAEAAGLGLKVLGDPDPPTNNYGVWEDEFKDLGLERCIEHAWKDTIVYLD 174
 DB 116 CGPAGLAAEAAGLGLKVLGDPDPPTNNYGVWEDEFKDLGLERCIEHAWKDTIVYLD 175
 QY 175 NDAPVLIGRAYGRVSRHLLHEELLKRCVSGVSYLDSKVRITERAGDGHSLVVCENEIFI 234
 DB 176 DDKPITIGRAYGRVSRHLLHEELLKRCVSGVSYLDSKVRITERAGDGHSLVVCENEIFI 235
 QY 235 PCRLATVASGAASGLLEVEYGVGGPRVCVQTAYGVEVEVENNPDPNLMVMDYRDMOQK 294
 DB 236 PCRLATVASGAASGLLQYEVGGPRVCVQTAYGVEVEVENNPDPNLMVMDYRDMOQK 295
 QY 295 LQCSSEYPTFLYVMPKTRLFEEETCLASKDAMPDGLLRKMSRLTKLGIGQVTKYE 354
 DB 296 VRSLEAEYPTFLYVMPKTRLFEEETCLASKDAMPDGLLRKMSRLTKLGIGQVTKYE 355
 QY 355 EWSYIPVGGSLPNTQKNLAFGAASMVHPATGYSVVRSLSEAPKYASVIAKILKQDNS 414
 DB 356 EWSYIPVGGSLPNTQKNLAFGAASMVHPATGYSVVRSLSEAPKYASVIAKILRETT 415
 QY 415 AYVVSQSSAVNISQMAWSSLPKRRQRAFFLGLGLIIVQDIEATRTFFRFLPT 474
 DB 416 KQINS-----NISQAWDTLWPPERKQRAFFLGLGLIIVQDIEATRTFFRFLPK 469
 QY 475 WNWGFLGSSLSFGLVLFMYMFLAPNSMRSLVRLHLLSDPQGVAMVRAYLE 528
 XX
 XX W09963055-A1.


```

XX SQ Sequence 516 AA;
Query Match 70.5%; Score 1934; DB 21; Length 516;
Best Local Similarity 71.4%; Pred. No. 1e-188;
Matches 365; Conservative 64; Mismatches 74; Indels 8; Gaps 2;

QY 22 NLSSSKLAYNTHRYGSSCRVDFQVRADGGSGSRSSVAYKEGVDFDEDFIKAGSSELLFVQ 81
DB 7 HMTATMAAFTCPWFMTSIRYTKQIKC---NAAKSOLVVKQETEEEDYKAGSSELLFVQ 63
QY 82 MQOTKSMKQAKLADKLPPPIPG-----ESVMDLVVIGCGPAGLSAAAKLGLKVLGLI 136
DB 64 MQONKSDMAQSLSQKLPVPVIGGGGSDNCILDLVVIGCGPAGLAGESAKLGLNVALI 123
QY 137 GPDLPFTNNYGWDEFFDLGLERCIHAWKDTIVYLDNDAPVILGRAYGRVSRHLLHEE 196
DB 124 GPDLPFTNNYGWDEFFDLGLERCIHAWKDTIVYLDNDAPVILGRAYGRVSRHLLHEE 183
QY 197 LKRCVSGSVYDLSKVERITEAGDGHSLVVCENEIFIPCLATVASGAASGLILEYEVG 256
DB 184 LLTRCMESGVSYLSKVERITEAPNGLSLIECEGNITIPCLATVASGAASGLILQYELG 243
QY 257 GPRVCVQTAYGVEVEVENNPDNLMVMDYRDYMQQKLCQSEEEYPTFLYVMPSPTRL 316
DB 244 GPRVCVQTAYGVEVEVESTPYDPSLMVMDYRDYTKHKSQSLEAQYPTFLYVMPSPTKV 303
QY 317 FFEETCLASKDAMPDPLLKRLKMSRLKTLGIVTKVYEEESYIPVGGSLPNTKOKNLAF 376
DB 304 FFEETCLASKDAMPPELLKTLKMSRLKTLGIRITKYTEEESYIPVGGSLPNTKOKNLAF 363
QY 377 GAAASVHPATGYSVVRSLSEAPKYASVIATKILKQDNSAYVYVSGOSSAVNTISMQAWSSLW 436
DB 364 GAAASVHPATGYSVVRSLSEAPKYAAVIAKILKQDNSAYVYVSGOSSAVNTISMQAWSSLW 423
QY 437 PKERRQRQRAFFLGLLELIVOLDIEATRTFFRFFRLPTWMMWGFLGSSLSFDDLVLFSMY 496
DB 424 PLERKQRQRAFFLGLLIVQMDIEGTRFFRFFRFLPTWMMWGFLGSSLSFDDLIFAFY 483
QY 497 MFVLAPNSMRSLVRHLLSDPSCGAVMVRAYL 527
DB 484 MFIIAPHSLRWGLVRHLLSDPTGGTMLKAYL 514

RESULT 11
AAAY54310
ID AAAY54310 standard; Protein; 516 AA.
XX AC AAAY54310;
XX DT 06-APR-2000 (first entry)
XX DE Amino acid sequence of a lycopene epsilon-cyclase protein.
KW Lycopene epsilon-cyclase; carotenoid biosynthesis pathway;
KW isopentenyl pyrophosphate isomerase; beta-carotene hydroxylase;
KW dimethylallyl pyrophosphate; transgenic plant; carotenoid content;
KW lethal photooxidative damage; colouring.
XX OS Tagetes erecta.
XX PN WO9963055-A1.
XX PD 09-DEC-1999.
XX PF 02-JUN-1999; 99WO-US12121.
XX PR 02-JUN-1998; 98US-0088724.
XX PR 02-JUN-1998; 98US-0088725.
XX PA (UYMA-) UNIV MARYLAND BALTIMORE.
XX PI Cunningham Fx, Sun Z;

```

```

XX DR WPI; 2000-116366/10.
XX PT New carotenoid biosynthesis genes used to provide transgenic plants
XX PI with altered carotenoid content -
XX PS Disclosure; Fig 25; 134pp; English.
XX CC The present sequence represents a lycopene epsilon-cyclase protein.
XX CC The enzyme is involved in the carotenoid biosynthesis pathway. The
XX CC specification also describes other enzymes involved in this pathway,
XX CC such as epsilon-cyclase, isopentenyl pyrophosphate (IPP) isomerase and
XX CC beta-carotene hydroxylase. IPP isomerase catalyses the reversible
XX CC conversion of IPP to dimethylallyl pyrophosphate (DMAPP). Beta-carotene
XX CC hydroxylase is responsible for hydroxylating the beta-endgroup in
XX CC carotenoids. Lycopene epsilon-cyclase is responsible for the formation
XX CC of epsilon-endgroups in carotenoids. The polynucleotides of the
XX CC invention are used to provide transgenic plants with altered
XX CC carotenoid content. These plants may, for example, may be protected
XX CC from lethal photooxidative damage or have altered colouring.
XX SQ Sequence 516 AA;
Query Match 70.5%; Score 1934; DB 21; Length 516;
Best Local Similarity 71.4%; Pred. No. 1e-188;
Matches 365; Conservative 64; Mismatches 74; Indels 8; Gaps 2;

QY 22 NLSSSKLAYNTHRYGSSCRVDFQVRADGGSGSRSSVAYKEGVDFDEDFIKAGSSELLFVQ 81
DB 7 HMTATMAAFTCPWFMTSIRYTKQIKC---NAAKSOLVVKQETEEEDYKAGSSELLFVQ 63
QY 82 MQOTKSMKQAKLADKLPPPIPG-----ESVMDLVVIGCGPAGLSAAAKLGLKVLGLI 136
DB 64 MQONKSDMAQSLSQKLPVPVIGGGGSDNCILDLVVIGCGPAGLAGESAKLGLNVALI 123
QY 137 GPDLPFTNNYGWDEFFDLGLERCIHAWKDTIVYLDNDAPVILGRAYGRVSRHLLHEE 196
DB 124 GPDLPFTNNYGWDEFFDLGLERCIHAWKDTIVYLDNDAPVILGRAYGRVSRHLLHEE 183
QY 197 LKRCVSGSVYDLSKVERITEAGDGHSLVVCENEIFIPCLATVASGAASGLILEYEVG 256
DB 184 LLTRCMESGVSYLSKVERITEAPNGLSLIECEGNITIPCLATVASGAASGLILQYELG 243
QY 257 GPRVCVQTAYGVEVEVENNPDNLMVMDYRDYMQQKLCQSEEEYPTFLYVMPSPTRL 316
DB 244 GPRVCVQTAYGVEVEVESTPYDPSLMVMDYRDYTKHKSQSLEAQYPTFLYVMPSPTKV 303
QY 317 FFEETCLASKDAMPDPLLKRLKMSRLKTLGIVTKVYEEESYIPVGGSLPNTKOKNLAF 376
DB 304 FFEETCLASKDAMPPELLKTLKMSRLKTLGIRITKYTEEESYIPVGGSLPNTKOKNLAF 363
QY 377 GAAASVHPATGYSVVRSLSEAPKYASVIATKILKQDNSAYVYVSGOSSAVNTISMQAWSSLW 436
DB 364 GAAASVHPATGYSVVRSLSEAPKYAAVIAKILKQDNSAYVYVSGOSSAVNTISMQAWSSLW 423
QY 437 PKERRQRQRAFFLGLLELIVOLDIEATRTFFRFFRLPTWMMWGFLGSSLSFDDLVLFSMY 496
DB 424 PLERKQRQRAFFLGLLIVQMDIEGTRFFRFFRFLPTWMMWGFLGSSLSFDDLIFAFY 483
QY 497 MFVLAPNSMRSLVRHLLSDPSCGAVMVRAYL 527
DB 484 MFIIAPHSLRWGLVRHLLSDPTGGTMLKAYL 514

RESULT 12
AAAY54308
ID AAAY54308 standard; Protein; 533 AA.
XX AC AAAY54308;
XX DT 06-APR-2000 (first entry)
XX DE Amino acid sequence of a lycopene epsilon-cyclase protein.

```

```

XX Lycopene epsilon-cyclase: carotenoid biosynthesis pathway;
KW isopentenyl pyrophosphate isomerase; beta-carotene hydroxylase;
KW dimethylallyl pyrophosphate; transgenic plant; carotenoid content;
XX lethal photooxidative damage; colouring.
XX
OS Lactuca sativa.
XX
PN W09963055-A1.
XX
PD 09-DEC-1999.
XX
XX 02-JUN-1999; 99WO-US12121.
XX
PR 02-JUN-1998; 98US-0088724.
PR 02-JUN-1998; 98US-0088725.
XX
PA (UYMA-) UNIV MARYLAND BALTIMORE.
XX
PI Cunningham FX, Sun Z;
XX
XX WPI; 2000-116366/10.
XX
XX New carotenoid biosynthesis genes used to provide transgenic plants
PT with altered carotenoid content
XX
XX Disclosure; Fig 25; 134pp; English.
XX
CC The present sequence represents a lycopene epsilon-cyclase protein.
CC The enzyme is involved in the carotenoid biosynthesis pathway. The
CC specification also describes other enzymes involved in this pathway,
CC such as epsilon-cyclase, isopentenyl pyrophosphate (IPP) isomerase and
CC beta-carotene hydroxylase. IPP isomerase catalyses the reversible
CC conversion of IPP to dimethylallyl pyrophosphate (DMAPP). Beta-carotene
CC hydroxylase is responsible for hydroxylating the beta-endgroup in
CC carotenoids. Lycopene epsilon-cyclase is responsible for the formation
CC of epsilon-endgroups in carotenoids. The polynucleotides of the
CC invention are used to provide transgenic plants with altered
CC carotenoid content. These plants may, for example, may be protected
CC from lethal photooxidative damage or have altered colouring.
XX
SQ Sequence 533 AA;
Query Match 70.5%; Score 1934; DB 21; Length 533;
Best Local Similarity 68.3%; Pred No. 1.1e-188;
Matches 375; Conservative 57; Mismatches 77; Indels 40; Gaps 5;
QY 1 MELLGVRNL-----ISSCPVMT-----FCTRNLSKSLAYNIHRYGSSCR 40
DB 1 MECFGARNMTATWVFTCPREFDCNIRHKFSLKQRRFTNLSASSSLRQIKCSAKSDRCV 60
QY 41 VDFQVRADGGSGSRSSVAYKEGF--VDEDFIKAGSGSELLFVOMQOTKSMKQAKLADKL 98
DB 61 VD-----KQISVADEEDYVKAGGSELFVQMORTKSMESQSKLSEKL 103
QY 99 PPIPGESVMDLVWIGCGPAGLSAAEAKGLKVLGPDLPFTNNYGVWDEDEKDLGL 158
DB 104 AQIPGNCILDLVWIGCGPAGLAAESAKGLNVLGPDLPFTNNYGVWQDEFGLGL 163
QY 159 ERCIEHANKDTIVYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVSGVSYLDSKVERITE 218
DB 164 EGCIEHSWKDTLVYLDNDADPIRGRAYGRVSRHLLHEELLKRCVSGVSYLSSKVERITE 223
QY 219 AGDGHSLVVCNEIPIPCRLATVAGSAAGKLEVEVGPRVCVOTAYGVEVEVNNPYD 278
DB 224 APNGYSLICEGNIIPICRLATVAGSAAGKLEVEVGPRVCVOTAYGVEVEVNNPYD 283
QY 279 PNLWFMFDYRDMQKQKQCESEYPTFLVMPMSPTLFFETCLASKDAMPDFLLKRL 338
DB 284 PDLWFMFDYRDSKHPESLEAKYPTFLVMAWSPKIFFEETCLASREAMPFNLLKSL 343
QY 339 MSRLKTLGIQVTKVYEEWSYIPVGGSLPNTQKNLAFGAAASWHPATGYSVVRSLSA 398
DB 344 MSRLKAMGIRITRTYEEWSYIPVGGSLPNTQKNLAFGAAASWHPATGYSVVRSLSA 403
QY 399 PNYASVIAKILQDINSAYVVGSSAVNISMOWSSLWPKERKQRAFLFGLLEIVQLD 458
DB 404 PNYAAVIAKILQDQSKEMIS-LGKYTNISKQAWETLWPLERKQRAFLFGLSHIVLMD 462
QY 459 IEATRTFFRTFFRLPTWMMWGLGSSLSGFDLVLFPSYMFVLAPNSMRKSLVRHLLSDPS 518
DB 463 LEGTRTFFRTFFRLPKWMMWGLGSSLSSTDLLIFALYMFVIAPIHSLRMLVRLHLLSDPT 522
QY 519 GAVMVRAYL 527
DB 523 GATWVKAYL 531
XX
RESULT 13
AAV52592
ID AAV52592 standard; Protein; 532 AA.
XX
AC AAV52592;
XX
DT 07-MAR-2000 (first entry)
XX
DE Romaine lettuce epsilon lycopene cyclase.
XX
KW Lycopene; epsilon cyclase; bicyclic epsilon carotene; biosynthesis;
KW carotenoid; epsilon ring; lutein; lutein; lutein; lutein; lutein;
KW transgenic plant; altered synthesis; overexpression;
KW nutritional value; pharmacology; colour; cancer; antisense;
KW accumulation; substrate; beta-carotene; vitamin A.
XX
OS Lactuca sativa.
XX
XX Key Location/Qualifiers
FT Misc-difference 36 /note= "Encoded by CGAAGA"
FT Misc-difference 461 /note= "Encoded by ATN"
XX
PN W09961399-A1.
XX
XX 02-DEC-1999.
XX
XX 25-MAY-1999; 99WO-US10461.
XX
XX 26-MAY-1998; 98US-0084222.
XX
XX (UYMA-) UNIV MARYLAND BALTIMORE.
XX
XX Cunningham FX;
XX
XX WPI; 2000-062667/05.
XX
XX N-PSDB; AAZ46328.
XX
XX New eukaryotic epsilon-lycopene cyclase and related DNA, used to
PT regulate carotenoid synthesis in plants and other hosts, e.g. for
PT treatment of cancer
XX
XX Claim 2; Fig 4; 38pp; English.
XX
XX This sequence represents epsilon lycopene cyclase from romaine
CC lettuce. This enzyme adds two epsilon rings to lycopene to
CC form bicyclic epsilon-carotene. Carotenoids with two epsilon rings
CC are found in significant amounts in relatively few plants (romaine
CC lettuce produces an abundance of such a carotenoid, lutein),
CC while carotenoids with two beta rings are ubiquitous and those
CC with one beta and one epsilon ring are common. The invention relates to
CC epsilon lycopene cyclases capable of forming bicyclic epsilon-carotene,
CC and also to methods of transforming known carotenoids into novel or rare
CC products. Epsilon lycopene cyclase nucleic acids may be used for
CC recombinant protein production and for the generation of transgenic
CC plants that have altered patterns of carotenoid synthesis. Such plants
CC may accumulate new or rare carotenoids or overexpress known carotenoids,

```

CC resulting in altered nutritional value, pharmacology or colour.
 CC Carotenoids isolated from such transgenic plants are useful
 CC therapeutically, e.g., for treatment of cancer. Alternatively, epsilon
 CC lycopene cyclase antisense nucleic acids can be used to cause
 CC accumulation of the substrates of the enzyme, specifically beta-carotene
 CC or derived products such as vitamin A. The isolated eukaryotic enzyme
 CC allows synthesis of epsilon, epsilon-carotene (which can be
 CC isolated from natural sources only in small amounts) on a large scale,
 CC enabling bulk production of derived carotenoids.

XX Sequence 532 AA;

Query Match 70.2%; Score 1925.5; DB 21; Length 532;
 Best Local Similarity 68.2%; Pred. No. 7.8e-188;
 Matches 374; Conservative 57; Mismatches 78; Indels 39; Gaps 5;

QY 1 MELLGVRNL-----ISSCPVWT-----FCTRNLSKSLAYNTHRYGSSCRV 41
 DB 1 MECFGARNWTAMVFTCPRTDCNIRHKFLLKQRTNLSASSSLRQIKCSAKSDRCV 60
 QY 42 DFOVRADGGSSRSVAYKEGF--VDEEDFIKAGSELFFVQMOTKSMERQAKLADKL 99
 DB 61 D-----KQGISVADEEDYKAGSELFFVQMOTKSMESQSLSEKLA 103
 QY 100 PIPFGESVMDLVWICGPGAGLSLAAEAKGLKVLGIDPLPFTNNYGVWEDEFKDLGLE 159
 DB 104 QIPGNCILDLWICGPGAGLAAESAKGLNVLGIDPLPFTNNYGVWODEFIGLLE 163
 QY 160 RCIEHAWKDTIVYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVSGVSYLDSKVERITEA 219
 DB 164 GCIEHSWKDTIVYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVSGVSYLDSKVERITEA 223
 QY 220 GDGSLVVCENEIPIPCRLATVASGAAGSKLLEYEVGPRVCVQTAYGVEVEVNNPYDP 279
 DB 224 PNGYSLICEGNITIPCLATVASGAAGSKLLEYEVGPRVCVQTAYGVEVEVNNPYDP 283
 QY 280 NLMVFMVDRYDWMQKQCSSEYPTFLYVMPMSPTLFFETCLASKDAMPFDLLKRLM 339
 DB 284 DLMVFMVDRYDWMQKQCSSEYPTFLYVMPMSPTLFFETCLASKDAMPFDLLKRLM 343
 QY 340 SRLTKLTGIVTKVTEEWYSYIPVGGSLPNTQKQLAFAGAAASVHPATGYSVRSLSRAP 399
 DB 344 SRLKAMGIRITTEEWYSYIPVGGSLPNTQKQLAFAGAAASVHPATGYSVRSLSRAP 403
 QY 400 KYASVIAKILKQDSAYVYVQSGSNAVNSMQAWSLWPKRKRORAFPLFGLLEIVOLDI 459
 DB 404 NYAAVIAKILRQDOSKEMIS-LGKYTNISKQAWETLWPLERKRORAFPLFGLSHIVLMDL 462
 QY 460 EATRTFTFTFRLPTWMMWGLGSSLSFSLDVLVFSMYMFLVAPNSMRMSLVRLHLSDPG 519
 DB 463 EGTRFTFTFRLPTWMMWGLGSSLSFSLDVLVFSMYMFLVAPNSMRMSLVRLHLSDPG 522
 QY 520 AVMVRAYL 527
 DB 523 ATMVKAYL 530

RESULT 14

AA54309
 ID AA54309 standard; Protein; 526 AA.

XX AC AA54309;

XX DT 06-APR-2000 (first entry)

XX DE Amino acid sequence of a lycopene epsilon-cyclase protein.

XX KW Lycopene epsilon-cyclase; carotenoid biosynthesis pathway;
 KW isopentenyl pyrophosphate isomerase; beta-carotene hydroxylase;
 KW dimethylallyl pyrophosphate; transgenic plant; carotenoid content;
 KW lethal photooxidative damage; colouring.

OS Lycopopersicon sp.

XX WO9963055-A1.
 PN 09-DEC-1999.
 PD 02-JUN-1999; 99WO-US12121.
 PF 02-JUN-1998; 98US-0088724.
 PR 02-JUN-1998; 98US-0088725.
 PS (UYMA-) UNIV MARYLAND BALTIMORE.
 XX Cunnigham FX, Sun 2;
 PI WPI; 2000-116366/10.
 DR New carotenoid biosynthesis genes used to provide transgenic plants
 PT with altered carotenoid content -
 XX Disclosure; Fig 25; 134pp; English.
 PS The present sequence represents a lycopene epsilon-cyclase protein.
 CC The enzyme is involved in the carotenoid biosynthesis pathway. The
 CC specification also describes other enzymes involved in this pathway,
 CC such as epsilon-cyclase, isopentenyl pyrophosphate (IPP) isomerase and
 CC beta-carotene hydroxylase. IPP isomerase catalyses the reversible
 CC conversion of IPP to dimethylallyl pyrophosphate (DMAPP). Beta-carotene
 CC hydroxylase is responsible for hydroxylating the beta-endgroup in
 CC carotenoids. Lycopene epsilon-cyclase is responsible for the formation
 CC of epsilon-endgroups in carotenoids. The polynucleotides of the
 CC invention are used to provide transgenic plants with altered
 CC carotenoid content. These plants may, for example, may be protected
 CC from lethal photooxidative damage or have altered colouring.

XX Sequence 526 AA;

Query Match 69.6%; Score 1909.5; DB 21; Length 526;
 Best Local Similarity 68.5%; Pred. No. 3.3e-186;
 Matches 368; Conservative 68; Mismatches 78; Indels 23; Gaps 6;

QY 1 MELLGVRNLISCPV-----WTFGTRNLSSK---LAYNTHRYGSSCRDVFQVRADGG 50
 DB 1 MECVGVQNVGAMAVLTPRLNRNSGG--ELCOEKSIFLAY--EQYESK-----NSS 48
 QY 51 SGRSSVAYKEGFVDEEDFIKAGSELFFVQMOTKSMERQAKLADKLPIPFSGSMDL 110
 DB 49 SGRSDCVYDKEDFADEEDYIKAGSQLVFMQOQKMDQCKSLDELRLQISAGQTVLDL 108
 QY 111 VVICGPAGLSLAAEAKGLKVLGIDPLPFTNNYGVWEDEFKDLGLERCIEHAWKDTI 170
 DB 109 VVICGPAGLSLAAEAKGLKVLGIDPLPFTNNYGVWEDEFKDLGLQACIEHVRDITI 168
 QY 171 VYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVSGVSYLDSKVERITEAGDGHSLVVCEN 230
 DB 169 VYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVSGVSYLDSKVERITEAGDGHSLVVCEN 228
 QY 231 EIFIPCLATVASGAAGSKLLEYEVGPRVCVQTAYGVEVEVNNPYDPNLMVFMVDRYD 290
 DB 229 DWIIPCRFVTVASGAAGSKLLEYEVGPRVCVQTAYGVEVEVNNPYDPNLMVFMVDRYD 288
 QY 291 MQKQCSSEYPTFLYVMPMSPTLFFETCLASKDAMPFDLLKRLHLSKLTGIVT 350
 DB 289 LRHDAQSLKAIPTFLYAMPSPTRVFETCLASKDAMPFDLLKRLHLSKLTGIVT 348
 QY 351 KYVEEWSYIPVGGSLPNTQKQLAFAGAAASVHPATGYSVRSLSRAPKASVIAKILK 410
 DB 349 EYVEEWSYIPVGGSLPNTQKQLAFAGAAASVHPATGYSVRSLSRAPKASVIAKILK 408
 QY 411 QDSAYVYVQSGSNAVNSMQAWSLWPKRKRORAFPLFGLLEIVOLDIETRTFTFTFF 470
 DB 409 QHYSKNMLT-SSSIPSTQAWNTLWPKRKRORAFPLFGLLEIVOLDIETRTFTFTFF 467
 QY 471 RLPTWMMWGLGSSLSFSLDVLVFSMYMFLVAPNSMRMSLVRLHLSDPGSAVMVRAYL 527

```
Db      468 RVPKVMQGFGLGSSLSADMLFAFYNEIIAPNDMRKGLIRHLLSDPGATLRTYL 524
```

RESULT 15

AAE09798
ID AAE09798 standard; Protein: 517 AA.

AAE09798;

DT 10-DEC-2001 (first entry)

Spinach lycopene epsilon cyclase protein.

Spinach; lycopene epsilon cyclase; lutein; therapeutic; alpha-carotene; nutritional supplement; food product.

KW nutritional supplement; food product.
KW nutritional supplement; food product.
KW nutritional supplement; food product.
KW nutritional supplement; food product.

OS **Spinacia oleracea.**

PN WO200166703-A1.

13-SEP-2001.

07-MAR-2001: 2001WO-US07178.

07-MAR-2000; 2000US-0187576.

AA (CRGI) CARGILL INC.

PI De Souza ML, Kollman SR, Schroeder WA:

WPI: 2001-582274/65.

Novel isolated lycopene epsilon cyclase polypeptide, useful for producing lutein, feed supplement and enhanced food products -

PS Claim 8; Fig 3; 37pp; English.

The invention relates to spinach lycopenes epsilon lipoxygenase protein and their corresponding cDNA molecule. Lycopenes epsilon lipoxygenase is useful for producing lutein in microorganisms. Lycopenes epsilon lipoxygenase is used as nutritional supplement or as sources of lutein for producing enhanced food products or other therapeutic purposes. The invention also relates to a method for producing lutein, where a single polypeptide converts alpha-carotene to lutein in the microorganism. The present sequence is spinach lycopenes epsilon lipoxygenase protein.

Sequence 517 AA:

Query Match	69.0%;	Score 1892;	DB 22;
Best Local Similarity	71.4%;	Pred. No. 2e-184;	Length 517;
Matches 355;	Conservative	61;	Mismatches 75;
			Indels 6;
			Gaps 1;

QY 37 SSCRVDFQVRADGGSGRSS-----VAYKEGFVDEEDFIKAGGSELLEFVOMOOTKSMEK 90

Db 19 SNVVECSRRSGSVLRANSDDSSCVIAPEDFANEEDFIKAGGSELLVYOMOONKAMDC 78

QY 91 QAKLADKLPPIPFGESVMDLVVIGCGPAGLSLAAEAAKGLKVGLIGPDLPTNNYGVWE 150

[illegible]

QY	151	DEFKDLGLERCIEHAWKDTIVYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVESGVSYLD	210
----	-----	--	-----

```

Db      139  DEFRA LGG CIEH VWRDTIVYIDDDNPIYIGRSYKVSROL LHKELVHRCL ESGVSYLN 198
      139  DEFRA LGG CIEH VWRDTIVYIDDDNPIYIGRSYKVSROL LHKELVHRCL ESGVSYLN 198

```

QY 211 SKVERITEAGDGHSLVVCENEIFIPCRLATVASGAASGKLLLEYEVGGPRVCVOTAYGVFV 270

Db 199 AKVENIMEGPDGHRVLVACERGV TIP CRLVTVASGAASGK LLEYEVGGPRVCVOTAYGVEV 258

QY 271 EVENNPYPNLMVFMDYRDYMQOKLOCSEEEYPTFLYVPMSPTRLFFETCLASKDAMP 330

Db 259 EVENSPYDPNVMVMDYRDYTKLSVQSLEAKYPTFLYAMPISPTRIFFFEETCLASVDAMP 318

THIS PAGE BLANK (USPTO)


```

QY 55 SSVAYKEGVDEDEPIKAGGSELLFVQMQQKSMKQAKLADKLPPIPFPGSVMDLVVIG 114
Db 56 SCVAVREDFADEEDEVKAGGSEILFVQMQQKMDQEQSKLDKLPPIGSGDGLDHHVIG 115
QY 115 CGPAGLSAAEAAKGLKVLGIGLDPPTNNYGVWEDEFKDLGLERCIEHAWKOTIVYLD 174
Db 116 CGPAGLAAEAAKGLKVLGIGLDPPTNNYGVWEDEFDGLQKCIHVVWRETIVYLD 175
QY 175 NDAPVLIGRAYGRYSRHLLEELKRCVSGSVYLDSEVERITEAGDGHSLVVCENEIFI 234
Db 176 DDKPITIGRAYGRYSRHLLEELKRCVSGSVYLDSEVERITEAGDGLRLVACDDNVI 235
QY 235 PCRLATVAGGAASKLLEYEVGGPRVCVQTAIGVEVEVENNPDPNLMVMDYRDMQOK 294
Db 236 PCRLATVAGGAASKLLEYEVGGPRVCVQTAIGVEVEVENNPDPNLMVMDYRDMQOK 295
QY 295 LOCSEEEYPTFLYVMPNPTLFEETCLASKDAMPDGLLKRKLMSRLKTLGIQVTKYE 354
Db 296 VRSLEAEYPTFLYAMPKTSRLFEETCLASKDAMPDGLLKRKLMSRLKTLGIQVTKYE 355
QY 355 EWSYIPVGGSLPNTQKNLAFGAAASMVHPATGYSVVRSLSSEAPKYASVIAEILRETT 414
Db 356 EWSYIPVGGSLPNTQKNLAFGAAASMVHPATGYSVVRSLSSEAPKYASVIAEILRETT 415
QY 415 AYVVGSSAVNISQAWSSLPKRRQRAFFLGLGLIIVOLDIEATRTFFRFLPT 474
Db 416 KOINS-----NISQAWDTLWPPERKQRAFFLGLGLIIVOLDIEATRTFFRFLPT 469
QY 475 WMMGFLGSSLSFDLVLFMYMFLAPNSMRSLVRLHLLSDPSGAVMVRAYLE 528
Db 470 WMMGFLGSLTSLGDLVLFALYMFVISPNNLRKGLNHLISDPTGATWIKYILK 523

```

RESULT 2

```

us-08-624-125-21
; Sequence 21, Application US/08624125
; Patent No. 574341
; GENERAL INFORMATION:
; APPLICANT: CUNNINGHAM JR., FRANCIS X.
; APPLICANT: SUN, ZAIEN
; TITLE OF INVENTION: GENES OF CAROTENOID BIOSYNTHESIS AND
; TITLE OF INVENTION: METABOLISM AND A SYSTEM FOR SCREENING SUCH GENES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/624,125
; FILING DATE: 29-MAR-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: KELBER, STEVEN B.
; REGISTRATION NUMBER: 30,073
; REFERENCE/DOCKET NUMBER: 2747-063-27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 524 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

```

```

; MOLECULE TYPE: protein
us-08-624-125-21
Query Match 71.4%; Score 1959.5; DB 1; Length 524;
Best Local Similarity 70.4%; Pred. No. 1e-200;
Matches 376; Conservative 60; Mismatches 81; Indels 17; Gaps 5;
QY 1 MELLGVRNL-----ISSCPVTFGTNRNLSSSLAYNIHRYGSSCRVDFQVRADGG--SGSR 54
Db 1 MECVGARFAAMAVSTFPSSWSCRKRPVVKVYSYRIRFG-LC-----SVRAGSGSSGSE 55
QY 55 SSVAYKEGVDEDEPIKAGGSELLFVQMQQKSMKQAKLADKLPPIPFPGSVMDLVVIG 114
Db 56 SCVAVREDFADEEDEVKAGGSEILFVQMQQKMDQEQSKLDKLPPIGSGDGLDHHVIG 115
QY 115 CGPAGLSAAEAAKGLKVLGIGLDPPTNNYGVWEDEFKDLGLERCIEHAWKOTIVYLD 174
Db 116 CGPAGLAAEAAKGLKVLGIGLDPPTNNYGVWEDEFDGLQKCIHVVWRETIVYLD 175
QY 175 NDAPVLIGRAYGRYSRHLLEELKRCVSGSVYLDSEVERITEAGDGHSLVVCENEIFI 234
Db 176 DDKPITIGRAYGRYSRHLLEELKRCVSGSVYLDSEVERITEAGDGLRLVACDDNVI 235
QY 235 PCRLATVAGGAASKLLEYEVGGPRVCVQTAIGVEVEVENNPDPNLMVMDYRDMQOK 294
Db 236 PCRLATVAGGAASKLLEYEVGGPRVCVQTAIGVEVEVENNPDPNLMVMDYRDMQOK 295
QY 295 LOCSEEEYPTFLYVMPNPTLFEETCLASKDAMPDGLLKRKLMSRLKTLGIQVTKYE 354
Db 296 VRSLEAEYPTFLYAMPKTSRLFEETCLASKDAMPDGLLKRKLMSRLKTLGIQVTKYE 355
QY 355 EWSYIPVGGSLPNTQKNLAFGAAASMVHPATGYSVVRSLSSEAPKYASVIAEILRETT 414
Db 356 EWSYIPVGGSLPNTQKNLAFGAAASMVHPATGYSVVRSLSSEAPKYASVIAEILRETT 415
QY 415 AYVVGSSAVNISQAWSSLPKRRQRAFFLGLGLIIVOLDIEATRTFFRFLPT 474
Db 416 KOINS-----NISQAWDTLWPPERKQRAFFLGLGLIIVOLDIEATRTFFRFLPT 469
QY 475 WMMGFLGSSLSFDLVLFMYMFLAPNSMRSLVRLHLLSDPSGAVMVRAYLE 528
Db 470 WMMGFLGSLTSLGDLVLFALYMFVISPNNLRKGLNHLISDPTGATWIKYILK 523

```

RESULT 3

```

us-09-201-641-6
; Sequence 6, Application US/09201641A
; Patent No. 6232530
; GENERAL INFORMATION:
; APPLICANT: Cunningham Jr, Francis X
; APPLICANT: Dellapenna, Dean
; TITLE OF INVENTION: Method for Regulating Carotenoid Biosynthesis in
; FILE REFERENCE: Quest 41-162
; CURRENT APPLICATION NUMBER: US/09/201,641A
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 516
; TYPE: PRT
; ORGANISM: Tagetes erecta
us-09-201-641-6
Query Match 70.5%; Score 1934; DB 4; Length 516;
Best Local Similarity 71.4%; Pred. No. 5.3e-198;
Matches 365; Conservative 64; Mismatches 74; Indels 8; Gaps 2;
QY 22 NLSSSKLAYNIHRYGSSCRVDFQVRADGGSSSVAYKEGVDEDEPIKAGGSELLFVQ 81
Db 7 HMTATMAAFTCPREMTSIRVTQKIKC---NAAKSQLVVKQIEEEDYVKKAGSELLFVQ 63
QY 82 MQQTKSMEKQAKLADKLPPIPFPG-----ESVMDLVVIGCGPAGLSAAEAAKGLKVL 136

```


Db 78 GRAQDFVILIIICAGPAGRLAEQVSKYGIKVCVDPDS-PLSMWPNNYGVWDFENLGLD 136

QY 161 CIEHAKDITVYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVSGVYLSKVERITEAG 220

Db 137 CLDHKPMTCVHINDNKTLYGRPYGRVSRKLLKLLNSCVENRVRFYKAKWKV-EHE 195

QY 221 DGHSLVVCNEIFPCRLATVASGAASKLLEYEVGGR-VCVQTAAYGVEVENNPYDP 279

Db 196 EFESSIVCDDGKKIRGLSVNDASGFAS-DFIEYD--KPRNHGYQIAHVLVEVDNHPFDL 252

QY 280 NLAYFMYRDY---MQOKLOCSEEEYPTFLYVMPMSPTLFFETCLASKDAMPFDLLKR 336

Db 253 DKMVLMDWRDHLGNPEYLRVNNAKEPTFLYAMPDFDRLVLEETSLVSRPVLSEMEVKR 312

QY 337 KMSRLKTLGQVTKVYEESYIPVGGSLPNTQKNLAFGAASAMVHPATGYSVVSLS 396

Db 313 RMVARLHLGKVRSVIEEKCVPIMGGLPRIPQNVMAIGNSGIVHPSTGYMVARSMA 372

QY 397 EAPKVASVIAKILQDNSAYVVGSSAVNISQAWSLPKKRRORAFPLFGLGLIVQ 456

Db 373 LAPVLAENIVGL---GSTRMIRGS---QLYHRVWNLWPLDRRCVRECYSGFMEYLLK 425

QY 457 LDIEATRTFFTRFLPMTWMMGFLGSSLSDFDLVLFMSYMFVLAPNSMRSLV 510

Db 426 LDLAGTRLDFAFDLPDKYWGQFLSSRLSVKELGLLSLCLFGLHGSNLTRLDIV 479

RESULT 6

US-09-134-607A-17

; Sequence 17, Application US/09134607A

; Patent No. 6252141

; GENERAL INFORMATION:

; APPLICANT: Joseph Hirschberg et al.

; TITLE OF INVENTION: POLYNUCLEOTIDES CONTROLLING THE EXPRESSION

; TITLE OF INVENTION: OF AND CODING FOR GENE B IN TOMATO AND USE

; TITLE OF INVENTION: OF SAME FOR ALTERING CAROTENOID

; TITLE OF INVENTION: BIOSYNTHESIS

; NUMBER OF SEQUENCES: 25

; CORRESPONDENCE ADDRESSES:

; ADDRESSEE: Mark M. Friedman c/o Anthony Castorina

; CITY: Arlington

; STATE: Virginia

; COUNTRY: United States of America

; ZIP: 22202

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk

; COMPUTER: Twinhead, Slimnote 890TX

; OPERATING SYSTEM: MS DOS version 6.2,

; SOFTWARE: Word for Windows version 2.0,

; CURRENT APPLICATION DATA:

; FILING DATE:

; CLASSIFICATION: 800

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Friedman, Mark M.

; REGISTRATION NUMBER: 33,883

; REFERENCE/DOCKET NUMBER: 325/12

; TELEPHONE: 972-3-5625553

; TELEFAX: 972-3-5625554

; TELEX:

; INFORMATION FOR SEQ ID NO: 17:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 498

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

US-09-134-607A-17

Query Match 26.2%; Score 719.5; DB 4; Length 498;

Best Local Similarity 40.1%; Pred. No. 4.7e-68;

Matches 164; Conservative 78; Mismatches 148; Indels 19; Gaps 9;

QY 109 DLVVIGCGPAGLSAABAAKGLKVLIGLIDPLPT---NNYGVWEDEKDLGLERCIEHA 165

Db 83 DVLIIAGPAGRLAEQVSKYGIKVCVDPDS-PLSMWPNNYGVWDFENLGLDCLDKH 141

QY 166 WKDTIVYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVSGVYLSKVERITEAGDGHSL 225

Db 142 WPMTCVHINDNKTLYGRPYGRVSRKLLKLLNSCVENRVRFYKAKWKV-EHEEFSS 200

QY 226 VVCNEIFPCRLATVASGAASKLLEYEVGGR-VCVQTAAYGVEVENNPYDPNLVME 284

Db 201 IVCDGKKIRGLSVNDASGFAS-DFIEYD--KPRNHGYQIAHVLVEVDNHPFDL 257

QY 285 MDYRDY---MQOKLOCSEEEYPTFLYVMPMSPTLFFETCLASKDAMPFDLLKRLMSR 341

Db 258 MDWRDHLGNPEYLRVNNAKEPTFLYAMPDFDRLVLEETSLVSRPVLSEMEVKRMV 317

QY 342 LKTLGQVTKVYEESYIPVGGSLPNTQKNLAFGAASAMVHPATGYSVVSLSSEAPKY 401

Db 318 LRHLGKVRSVIEEKCVPIMGGLPRIPQNVMAIGNSGIVHPSTGYMVARSMA 377

QY 402 ASVTAKILQDNSAYVVGSSAVNISQAWSLPKKRRORAFPLFGLGLIVOLDIEA 461

Db 378 AEAIVEGL---GSTRMIRGS---QLYHRVWNLWPLDRRCVRECYSGFMEYLLKDL 430

QY 462 TRTFRTRFFTRFLPMTWMMGFLGSSLSDFDLVLFMSYMFVLAPNSMRSLV 510

Db 431 TRRLDFAFDLPDKYWGQFLSSRLSVKELGLLSLCLFGLHGSNLTRLDIV 479

RESULT 7

US-09-134-607A-18

; Sequence 18, Application US/09134607A

; Patent No. 6252141

; GENERAL INFORMATION:

; APPLICANT: Joseph Hirschberg et al.

; TITLE OF INVENTION: POLYNUCLEOTIDES CONTROLLING THE EXPRESSION

; TITLE OF INVENTION: OF AND CODING FOR GENE B IN TOMATO AND USE

; TITLE OF INVENTION: OF SAME FOR ALTERING CAROTENOID

; TITLE OF INVENTION: BIOSYNTHESIS

; NUMBER OF SEQUENCES: 25

; CORRESPONDENCE ADDRESSES:

; ADDRESSEE: Mark M. Friedman c/o Anthony Castorina

; STREET: 2001 Jefferson Davis Highway, Suite 207

; CITY: Arlington

; STATE: Virginia

; COUNTRY: United States of America

; ZIP: 22202

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk

; COMPUTER: Twinhead, Slimnote 890TX

; OPERATING SYSTEM: MS DOS version 6.2,

; SOFTWARE: Word for Windows version 2.0,

; CURRENT APPLICATION DATA:

; FILING DATE:

; CLASSIFICATION: 800

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Friedman, Mark M.

; REGISTRATION NUMBER: 33,883

; REFERENCE/DOCKET NUMBER: 325/12

; TELEPHONE: 972-3-5625553

; TELEFAX: 972-3-5625554

; TELEX:


```
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/624,125
;; FILING DATE: 29-MAR-1996
;; CLASSIFICATION:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: KELBER, STEVEN B.
;; REGISTRATION NUMBER: 30,073
;; REFERENCE/DOCKET NUMBER: 2747-063-27
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 703-413-3000
;; TELEFAX: 703-413-2220
;; INFORMATION FOR SEQ ID NO: 20:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 456 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-624-125-20

Query Match 25.6%; Score 701.5; DB 1; Length 456;
Best Local Similarity 36.0%; Pred. No. 3.4e-66;
Matches 169; Conservative 96; Mismatches 161; Indels 43; Gaps 17;

QY 62 GPVDEEDFTKAGSGELLFVQMOTKSM---EKQAKLADKLPIPFPGESVMDLVVIGCGPA 118
DB 16 GFV--KSFCKFCGCVCKSSALLELPETKKNLDFELPMYDFSGKGVDLAVVGGGPA 73
QY 119 GLSLAAEAAGLQKLVGLIG-PDLPTNNYGVWEDEFKDLGLERCIEHAWKDTIYVLDNDA 177
DB 74 GLAVAQVSEAGLSVCSIDPPKLIWPNNYGVWVDEFEAMDLLDCLDATWSGA-VYIDDTK 132
QY 178 PVLIGRAYGRVSRHLLHEELKRCVSGSYLDSKVERITEAGDGHSLVVCNEEIFPCR 237
DB 133 DL---RPGRVNRKQKSKMOKCI-NGVKFQAKVIRVIH--EKSMLIC-NDGTIQAT 185
QY 238 LATVASGAAGSKLLEYVG-GRVCVQTAYGVEVEENPNYDNLVFMVMDYRD-YMOOKL 295
DB 186 VVLDTATGS--RLVQYDKPNPG--YQVAYGILAEVEEHPDK--WVFMDWRDLSHNL 239
QY 296 QCSSEYPTFLVMPKSPTRLFEEETCLASKDAMPDFLLKRLMSRLKTLGIQVTKVVEE 355
DB 240 K-ERNISPTFLVAMPSSNRIFLEETSLVARPLRMDDOIQRNVARLH-LGIKVKSI 297
QY 356 EWSYIPVGSGLPNTQKNLAFGAAASMVHPATGYSVRSLSAPKYASVIAKILKQDNSA 415
DB 298 EHCVIPMGSLPLVLPQRVVIGGTAGMVHPSTGYMVARTLAAAPVVANAI----- 347
QY 416 YVVGQSSAVNISMQAWSLWPKRKQRAFLGLELIVQLDIEATRTFFTFPLPTW 475
DB 348 -IYLGSESGELSAEYWKDLWPIERRQRREFFCFGMDILLKLDLPATRRFFDAFFDLEPR 406
QY 476 MWGFLGSSLSFDLVLFMYMFVLAPNSMR-----MSLVRLHLS 516
DB 407 YHGFLLSRLLFPELIVFGLSLFASHASNTSRINTKGYPLVMINLLQD 455

RESULT 10
US-08-399-561-2
; Sequence 2, Application US/08399561
; Patent No. 5792903
; GENERAL INFORMATION:
; APPLICANT: Hirschberg, Joseph
; APPLICANT: Cunningham Jr., Francis X.
; APPLICANT: Gantt, Elisabeth
; TITLE OF INVENTION: Lycopen Cyclyase Gene
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kohn & Associates
; STREET: 30500 No. 5792903thwestern Highway, Suite 410
; CITY: Farmington Hills
; STATE: MI
; COUNTRY: US
```

```
;; ZIP: 48334
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/399,561
;; FILING DATE:
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Kohn, Kenneth I.
;; REGISTRATION NUMBER: 30,955
;; REFERENCE/DOCKET NUMBER: P-305 (Hebrew Univ.)
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 810-539-5050
;; TELEFAX: 810-539-5055
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 411 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-399-561-2

Query Match 22.1%; Score 605; DB 1; Length 411;
Best Local Similarity 35.8%; Pred. No. 6.1e-56;
Matches 151; Conservative 60; Mismatches 165; Indels 46; Gaps 10;

QY 107 VMDLVVICGAGLSLAAEAAGLGLKV-GL--IGDLPFTNNYGVWEDEFKDLGLERCIE 163
DB 1 VFDALVIGSGPAGLAAIAAELAQRLKQGLSPVDFPHWENTYGIWGPDELGLSLGHLFG 60
QY 164 HAWKDTIYVLDNDAPVLIGRAYGRVSRHLLHEELKRCVSGSYLDSKVERITEAGDGH 223
DB 61 HRWSCVSYF-GEAPVQHOYNYGLFDRALQLOOHNLRCQEGGLQOLGKAAAI--AHDSH 117
QY 224 SLWVCENEIFPCRATVASGAASKLLEYEVGGPR-----VCVQYAYGVEVEV 272
DB 118 HSCV-----TTAAGQELQARLVVDVTTGHAAFIQPHSDAIAAYAAAYGIGQF 165
QY 273 ENPNYDNLVMDYR-DYMOOKLOCSEEEY---PTELYVMPMSPTRLFEEETCLASKDA 328
DB 166 SPPPIEHOFLVMDYRSHL-----SPEEROLPTFLFAMDLGNDVTFVEETSLAACP 219
QY 329 MPFDLLKRLKSLRKLTLGIQVTKVYEESYIIPVGSGLPNTQKNLAFGAAASMVHPATG 388
DB 220 IPYDLKQRLYQRLATRGVTQVIQHEEYCLFPMNLPLPDLTQSVVGVGGAASMVHPASG 279
QY 389 YSVVRSLSAPKYASVIAKILKQDNSAYVVGQSSAVNISMQAWSLWPKRKQRAFL 448
DB 280 YMVGALLRRAPDLANAIAAGLNASSL-----TTAELATQAWRLMPTKIRKHYIQ 332
QY 449 FGLELIVOLDIEATRTFFTFPLPTWMMWGLGSSLSFDLVLFMYMFVLAPNSMRMS 508
DB 333 FGLEKLMRFSEAQLNHHFQTFGLPKEQWYGLTNTLSLPELIQAMRLRFAQAPNDVRWG 392
QY 509 LV 510
DB 393 LM 394

RESULT 11
US-09-147-009-11
; Sequence 11, Application US/09147009
; Patent No. 6153815
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Process for Raising Squalene Levels in Plants
; TITLE OF INVENTION: and DNA Sequences Used Therefor
; NUMBER OF SEQUENCES: 11
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```
;
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/147,009
; FILING DATE:
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 503 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-09-147-009-11

Query Match 5.1%; Score 139; DB 4; Length 503;
Best Local Similarity 23.2%; Pred. No. 6.3e-06;
Matches 124; Conservative 66; Mismatches 184; Indels 160; Gaps 27;

QY 89 EKQAKLADKLPPIPGSGVMDLVVIGCGPAGLSLAABAAKLGKVLIGPDL--PFTNNY 146
;| | | | | : : : : | | | | | : : : : |
Db 30 KKVAKLPDAATEVR-RDGDADVIIVGAGVGSALAYALAKDGRVHVIERDMREPVR-- 85
;| | | | | : : : : | | | | | : : : : |

QY 147 GWNEDEF-----KDLGLERCIE-----HAWKDTIVYL-----DNDAPV-LI 181
;| | | | | : : : : | | | | | : : : : |

Db 86 --MMGEFMQCGRLLLSKLGLEDCEIGDIQIATGLAYKDGQKALVSFPEDNDFPEPT 143
;| | | | | : : : : | | | | | : : : : |

QY 182 GRAY--GRVSRHL-----LHEELKRCVE-----SGVSYLDSKVERITEAGDGH 223
;| | | | | : : : : | | | | | : : : : |

Db 144 GRAFYNGRFVQRLRQKASSLPTVQLEEGTVKSLIEKGVIKGYTKNS-----AGE-- 194
;| | | | | : : : : | | | | | : : : : |

QY 224 SLVVCENEIIPCKLATVAGSAAS-----GKLLYEYVGGPRVCVOTAYGVEVE 273
;| | | | | : : : : | | | | | : : : : |

Db 195 -----ETTAFAP--LTVCDCGYSNLRSSVNDNNAEIVSYQVG-----YVSKNCQLE-- 239
;| | | | | : : : : | | | | | : : : : |

QY 274 NNPDNPLMVFMDYRDYMQKLCSEEEYPTFLYVPMSPSTRLEFEETCLASDKAM---- 329
;| | | | | : : : : | | | | | : : : : |

Db 240 -----DPEKLKLMKSPFTMLYQISSYDVRVCMEIFPGNIPISNGEMAYLKNMTAPQV 295
;| | | | | : : : : | | | | | : : : : |

QY 330 PFDLLKRLMSRLKTLGIQV-----TKVYEWEWSYIPVGGSLPNTPEQKNLAFGAASMVHP 385
;| | | | | : : : : | | | | | : : : : |

Db 296 PPEL--RKIFLKGIDEAQIKAMPTKMEATLS-----EKQGVIVLGDFAFNHRP 343
;| | | | | : : : : | | | | | : : : : |

QY 386 --ATGYSV-----RSLSEAPKYASVIAKILKQDMSAYV--SGOSSAYNIS 428
;| | | | | : : : : | | | | | : : : : |

Db 344 AIASGMVVLSDILILRRLQLPRLNLSANKVSEVI-----KSFYVIRKPSATVNTL 396
;| | | | | : : : : | | | | | : : : : |

QY 429 MQAWSSLW-----PKERKQRQRAFFLFGLELIVQLDIEATRTFTTFFRL-----P 473
;| | | | | : : : : | | | | | : : : : |

Db 397 GNAFSQVLIASDTDAKAMRQGCDFYL-----SSGGFRTSGMALLGGMNRP 444
;| | | | | : : : : | | | | | : : : : |

QY 474 TMMWGFGLSGSLSFDLVLFSTMYMFLAPNSMRSLVRLHLSLDPGAVMVRAYL 527
;| | | | | : : : : | | | | | : : : : |

Db 445 LSLIFLHCGITLSSIGQLLSPFPSPGLGIWHSRLFGVSYQMLSPAYAAAYRKSVM 498
;| | | | | : : : : | | | | | : : : : |

RESULT 12
US-08-095-726-14
; Sequence 14, Application US/08095726
; Patent No. 5530188
; GENERAL INFORMATION:
; APPLICANT: Ausich, Rodney L.
; APPLICANT: Brinkhaus, Friedhelm L.
; APPLICANT: Mukharji, Indrani
; APPLICANT: Proffitt, John H.
; APPLICANT: Yarger, James G.
; APPLICANT: Yen, Huel-Che B.
; TITLE OF INVENTION: Beta-Carotene Biosynthesis in
; NUMBER OF SEQUENCES: 79
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Amoco Corp., Patents and Licensing Dept
; STREET: 200 E Randolph St
; CITY: Chicago

; STATE: IL
; COUNTRY: USA
; ZIP: 60680-0703
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/095,726
; FILING DATE: 21-JUL-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/785,566
; FILING DATE: 30-OCT-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Galloway, No. 5530188val B
; TELEPHONE: 3128567180
; TELEFAX: 3128564972
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 374 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-095-726-14

Query Match 4.9%; Score 134.5; DB 1; Length 374;
Best Local Similarity 20.8%; Pred. No. 1.2e-05;
Matches 87; Conservative 61; Mismatches 167; Indels 103; Gaps 16;

QY 109 DLVVIGCGPA-GL---SLAAEAAKLGKVLIGPDLPTNNYGVWEDEF---KDLGLERC 161
;| | | | | : : : : | | | | | : : : : |

Db 3 DLIVVGGLANGLLAWLRQRYPOLNLLIEAGEQPGNHNWSHEDDLTPGQAWLAPL 62
;| | | | | : : : : | | | | | : : : : |

QY 162 IEHAWKDTIVYLDNDAPVLIGRAYGRVSRHLLHEELKRCVESGVSYLDSKVERITEA-- 219
;| | | | | : : : : | | | | | : : : : |

Db 63 VAHAWPGVEYQFPD-----LRRRLARG--YYSITSERFAEALH 98
;| | | | | : : : : | | | | | : : : : |

QY 220 ---CDGSHLVVCENEIIPCELATV-----ASGAA--SKLLEYEVGGPRVCVOTAYG 267
;| | | | | : : : : | | | | | : : : : |

Db 99 QALGE-----NIWLNCVSEVLPNSVRLANGEALLAGAVIDGRVTVASSAMOTGYQ 149
;| | | | | : : : : | | | | | : : : : |

QY 268 VEVEVE---NNPDNPLMVFMDYRDYMQKLCSEEEYPTFLYVPMSPSTRLEFEETCLA 324
;| | | | | : : : : | | | | | : : : : |

Db 150 LFLCQQWRLTQPHGLTPILMDATVAQQGYR-----FVYTLPLSADTLLEDTRYA 201
;| | | | | : : : : | | | | | : : : : |

QY 325 SKDAMPEDLLKRLKMSRLKTLGIQVTKVYEWEWSYIP-----VGGSLPNTPEQKNLAFGAAA 380
;| | | | | : : : : | | | | | : : : : |

Db 202 NVPQRDDNALRQTVDYAHSGKWLQALERETGCLPTWRTVSRLCGPMRRRAASGMRA 261
;| | | | | : : : : | | | | | : : : : |

QY 381 SMVHPATGYS-----VVRSLSEAPKYASVIAKILKQDMSAYVYVGGSSAVNISMQAWSS 434
;| | | | | : : : : | | | | | : : : : |

Db 262 GLFHPHTTGYSLPLAVALADAIADSPRLGSV-----PLYQLTRQFAER 303
;| | | | | : : : : | | | | | : : : : |

QY 435 LWPKE---RKRQRQRAFFLFGLELIVQLDIEATRTFTTFFRLPTWMMWGFGLSGSLSF 489
;| | | | | : : : : | | | | | : : : : |

Db 304 HWRQGFRLNLRMLFLAGRE-----ENRWRYMQRYFGLPEPTVERFYAGRLSLFD 354
;| | | | | : : : : | | | | | : : : : |

RESULT 13
US-08-096-623A-14
; Sequence 14, Application US/08096623A
; Patent No. 5684238
; GENERAL INFORMATION:
; APPLICANT: Ausich, Rodney L.
; APPLICANT: Brinkhaus, Friedhelm L.
; APPLICANT: Mukharji, Indrani
; APPLICANT: Proffitt, John H.
; APPLICANT: Yarger, James G.
; APPLICANT: Yen, Huel-Che B.
; TITLE OF INVENTION: Biosynthesis of zeaxanthin and
```

TITLE OF INVENTION: Glycosylated Zeaxanthin in Genetically Engineered Hosts
NUMBER OF SEQUENCES: 104
CORRESPONDENCE ADDRESS:

ADDRESSEE: Welsh & Katz, Ltd.

STREET: 120 S. Riverside Plaza, 22nd Floor

CITY: Chicago

STATE: IL

COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/096,623A

FILING DATE: 22-JUL-1993

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/805,061

FILING DATE: 09-DEC-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/662,921

FILING DATE: 28-FEB-1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/562,674

FILING DATE: 03-AUG-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/525,551

FILING DATE: 18-MAY-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/487,613

FILING DATE: 02-MAR-1990

ATTORNEY/AGENT INFORMATION:

NAME: Gamson, Edward P.

REGISTRATION NUMBER: 29,381

REFERENCE/DOCKET NUMBER: AMO-006.1

TELEPHONE: (312) 653-1500

TELEFAX: (312) 655-1501

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 374 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-096-623A-14

Query Match 4.9%; Score 134.5; DB 1; Length 374;
Best Local Similarity 20.8%; Pred. No. 1.2e-05;
Matches 87; Conservative 61; Mismatches 167; Indels 103; Gaps 16;

QY 109 DLVWIGCGPA-GL---SLAAEAAKGLKVLGIPDLPTNNYGVWDEF---KDLGLERC 161

Db 3 DLIVGGGLANGLIANRLRQYQPNLLLIAGEPGNGNHTWSPHEDDLTPGQHAWLAPL 62

QY 162 IEHAWKOTIVYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVSGVSYLDSKVERITEA-- 219

Db 63 VAHAWPGYEVQFPD-----LRRRLARG--YYSITSERFAEALH 98

QY 220 ---GDGHSLLVCENEIFIPCRLATV-----ASGAA--SGKLLYEYVGGRVPCVQTAYG 267

Db 99 QALGE-----NWLNCVSEVLPNSVRLANGEALLAGAVIDGRGVTAASSAMQTGYQ 149

QY 268 VEVEVE---NNPYDPNLMVMDYDMQKLCQSEEEYPTLYVMPSPTRLFEETCLA 324

Db 150 LFLGQOWRLTQPHGLTVPILMDAIVAAQQGYR-----FVYILPLSADTLLIEDIYIA 201

QY 325 SKDAMPFOLLKKRLMSRLTKTIGIQVTKYEEWSYIP-----VGGSLPNTQKNLAFGAAA 380

Db 202 NVPQRDDNALQRTVTDYAHSGKWLAQLEREETGCLPTITWRVTSRLCGPMRRRAASGMRA 261

QY 381 SMVHPATGYS-----VVSRLSEAPKYASVIKILKQDNSAVVSGQSSAVNISMQAWSS 434
Db 262 GLFHPTTGYSLPLAYALADAIADSPELGSV-----PLYQLTRFAER 303
QY 435 LMPKE---RKORAFPLFGLLEIVQLDIEATRTFTFFRFLPTMMWGLGSSLSFD 489
Db 304 HWRQGFRLNMLFLAGRE-----ENRWVRVMQRFTYGLPEPTVERFYAGRLSLFD 354

RESULT 14

US-08-663-310-6

; Sequence 6, Application US/08663310

; Patent No. 5811273

; GENERAL INFORMATION:

; APPLICANT: MISA, No. 5811273ihiko

; APPLICANT: KONDO, Keiji

; APPLICANT: KAJIWARA, Susumu

; APPLICANT: YOKOYAMA, Akihiko

; TITLE OF INVENTION: DNA STRANDS USEFUL FOR THE SYNTHESIS OF.

; TITLE OF INVENTION: XANTHOPHYLLS AND THE PROCESS FOR PRODUCING THE

; TITLE OF INVENTION: XANTHOPHYLLS

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Foley & Lardner

; STREET: 3000 K Street, N.W., Suite 500

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20007-5109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/663,310

; FILING DATE: 23-SEP-1996

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: WO PCT/JP94/02220

; FILING DATE: 26-DEC-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 6-235917

; FILING DATE: 05-SEP-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 5-348737

; FILING DATE: 27-DEC-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: BENT, Stephen A.

; REGISTRATION NUMBER: 29,768

; REFERENCE/DOCKET NUMBER: 49441/109

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202)672-5300

; TELEFAX: (202)672-5399

; TELEX: 904136

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 386 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-663-310-6

Query Match

Best Local Similarity 4.7%; Score 129; DB 2; Length 386;

Matches 92; Conservative 56; Mismatches 165; Indels 108; Gaps 18;

QY 109 DLVWIGCGPAG--LSLAAEAAKGLKVLGLI---GPDLPFTNNYGVWDEFKDLG----- 157

Db 4 DVLVLAGLANGLIATLALRAAPDLRLVLLDHAAGPSDGH-----WSCHDPDLSPDLA 58

QY 158 -LERCIEHAWKDTIVYLDNDAPVLIGRAYGRVSRHLLHEELLKRCVE-----SGVSYLDSK 212

Db 59 RLKPLRRANWPDQEVFRPHARRL-ATGYGSLDGAALADAVVRSGAEIRWDSIALDLDAQ 117
QY 213 VERIT-----EAGDGHSLVVCNEIFIPCLRLATVASGAASGKLLLEYEVGPRVCVQTAYG 267
Db 118 GATLSCGTTRIEAG-----AVLDGRGAQPSRHLLTVG-----FQKFFVG 153
QY 268 VEVEVENNPDNLVFMVMDYRDYMOQKLCSEEEYPTFLYVMPMSPTLFFETCLASKD 327
Db 154 VEIETDRPHGVPRPMI-MD-----AVTQODGYRFYLLFPFSTRILLIEDTRYSDGG 204
QY 328 AMPFDLLKRLKMSRLTKTLGIQVTKVYEEWSPV-----GGSLPNTPEQKN 373
Db 205 DLDDDLAALAAASHDYARQCGWTGAEVRRER-GILPIALAHDAAGFWADHAAGPVP-----257
QY 374 LAFGAASAMVHPATGYSVVRSLSSEAPKYASVIKILKQDNSAVVYVSGQS-----SAVNIS 428
Db 258 --VGLRAGFFHVTGYSL-----PYAAQVAD-----VWAGLSGPGTDLALRGA 298
QY 429 MOAWSSLMWPKERKORAFLEGLLEIVOLDIEATRTFFTRFLPTMMWMMGLGSSLSF 488
Db 299 IRDYAI---DRARRDRFLRLNMLFRGCAPDRRYTLLQRFYRMPHGLIERFYAGRLSVA 355
QY 489 D 489
Db 356 D 356

RESULT 15
US-09-006-491-6
; Sequence 6, Application US/09006491
; Patent No. 5972690
; GENERAL INFORMATION:
; APPLICANT: MISA, No. 5972690ihiko
; APPLICANT: KONDO, Keiji
; APPLICANT: KAJIWARA, Susumu
; APPLICANT: YOKOYAMA, Akihiro
; TITLE OF INVENTION: DNA STRANDS USEFUL FOR THE SYNTHESIS OF
; TITLE OF INVENTION: XANTHOPHYLLS AND THE PROCESS FOR PRODUCING THE
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/006,491
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/663,310
; FILING DATE: 23-SEP-1996
; APPLICATION NUMBER: WO PCT/JP94/02220
; FILING DATE: 26-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 6-235917
; FILING DATE: 05-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 5-348737
; FILING DATE: 27-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 49441/109
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300

TELEFAX: (202)672-5399
TELEX: 904136
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 386 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-006-491-6

Query Match 4.7%; Score 129; DB 2; Length 386;
Best Local Similarity 21.9%; Pred. No. 4.7e-05;
Matches 92; Conservative 56; Mismatches 165; Indels 108; Gaps 18;

QY 109 DLVITGCGPAG--LSLAAEAAGKLGKVGLI-----GPDLPFTNNYGVWEDEFKDLG-----157
Db 4 DVLLAGAGLANGLIALALRAAPDLRVLLDHAAGPSDGH-----WSCHDDPLSDPDLA 58
QY 158 -LERCIEHAWKDTIVYLDNDAPVLIIGRAYGRVSRHLLHEELKRCVE-----SGVSYLDSK 212
Db 59 RLKPLRRANWPDQEVFRPHARRL-ATGYGSLDGAALADAVVRSGAEIRWDSIALDLDAQ 117
QY 213 VERIT-----EAGDGHSLVVCNEIFIPCLRLATVASGAASGKLLLEYEVGPRVCVQTAYG 267
Db 118 GATLSCGTTRIEAG-----AVLDGRGAQPSRHLLTVG-----FQKFFVG 153
QY 268 VEVEVENNPDNLVFMVMDYRDYMOQKLCSEEEYPTFLYVMPMSPTLFFETCLASKD 327
Db 154 VEIETDRPHGVPRPMI-MD-----AVTQODGYRFYLLFPFSTRILLIEDTRYSDGG 204
QY 328 AMPFDLLKRLKMSRLTKTLGIQVTKVYEEWSPV-----GGSLPNTPEQKN 373
Db 205 DLDDDLAALAAASHDYARQCGWTGAEVRRER-GILPIALAHDAAGFWADHAAGPVP-----257
QY 374 LAFGAASAMVHPATGYSVVRSLSSEAPKYASVIKILKQDNSAVVYVSGQS-----SAVNIS 428
Db 258 --VGLRAGFFHVTGYSL-----PYAAQVAD-----VWAGLSGPGTDLALRGA 298
QY 429 MOAWSSLMWPKERKORAFLEGLLEIVOLDIEATRTFFTRFLPTMMWMMGLGSSLSF 488
Db 299 IRDYAI---DRARRDRFLRLNMLFRGCAPDRRYTLLQRFYRMPHGLIERFYAGRLSVA 355
QY 489 D 489
Db 356 D 356

Search completed: May 21, 2003, 22:12:39
Job time : 28 secs

THIS PAGE BLANK (USPTO)